2011

The Impossibility of Agnostic Discrimination Law

Deborah M. Weiss

Available at: https://works.bepress.com/deborah_weiss/10/
The Impossibility of Agnostic Discrimination Law

Deborah M. Weiss*

Abstract

The extent to which discrimination continues to play a role in American life remains a deeply contentious issue. Yet it seems irrelevant to the typical individual employment discrimination case, which apparently raises a narrow factual question: did a particular employer treat a particular employee differently because of membership in a protected group?

In this Article I argue that the trier of fact in virtually every discrimination case must use background assumptions about the societal pattern of discrimination. Discrimination law cannot be agnostic about the nature and frequency of discrimination. It can only determine how and by whom such assessments are made.

For many years, the Supreme Court introduced background assumptions through burdens of proof to govern findings of fact. It has now retreated from this role, forcing triers of fact to rely on their own background assumptions and creating dismayingly inconsistent case outcomes.

In recent years, expert witnesses have provided so-called social framework evidence about societal discrimination patterns in a specialized set of class actions including the landmark Dukes v. Wal-Mart which the Supreme Court will decide this year. Increased use of social framework evidence may reduce the role played by the personal views of judges and juries. On the other hand, considerations of both consistency and judicial economy suggest the undesirability of turning each and every discrimination trial into a forum for examining the pattern of discrimination in the United States today. To avoid this problem I propose an agenda for incorporating social frameworks into discrimination doctrine.

17 June 2011
Forthcoming Utah Law Review, Fall 2011

* © 2011 Deborah M. Weiss. Visiting Scholar, McCombs School of Business, University of Texas-Austin, Deborah.Weiss@mccombs.utexas.edu. This paper has been greatly improved by the comments of Bernie Black, Kingsley Browne, Jack Getman, Alvin Goldman, Jay Koehler, Dick Markovits, Gregory Mitchell, John Monahan, Paul Secunda, Laurens Walker, as well as those of participants in seminars at the University of Texas-Austin School of Law and the University of Texas-Austin McCombs School of Business.
I. INTRODUCTION

The extent and nature of discrimination in American life remains a deeply contentious issue. Judicial background assumptions about the societal pattern of discrimination—its manifestations, causes, and frequency—play a key role in the resolution of many doctrinal issues, and every Supreme Court nominee’s opinions on this matter are closely scrutinized.

Yet only a small set of discrimination cases depend on substantive doctrinal issues. The typical employment discrimination case seems to raise a simple factual question: did a particular employer treat a particular employee differently because of membership in a protected group? The larger societal pattern of discrimination, untethered to the facts of that dispute, might appear to fail the threshold “relevance” requirement of evidence law. Even if evidence of the societal pattern of discrimination were relevant, it might seem substantially more prejudicial than probative, since no individual defendant should be held responsible for the wrongs of society as a whole.

In this Article, I argue that whenever the factual occurrence of differential treatment is at issue, triers of fact must make background assumptions about the societal pattern of discrimination. The role of these background assumptions can be seen by considering how we would assess two different claims of discrimination. Consider a Southern factory in 1965. A black worker produces enough evidence to convince us that discrimination has occurred. Now suppose that a white worker at the same plant produces the same evidence and claims reverse discrimination. We would—and rationally should—be far more reluctant to draw an inference of discrimination in light of our knowledge about the relative likelihood of discrimination against blacks and against whites in that time and place. For example, if both plaintiffs produced evidence that comparable employees of another race were treated better, we would scrutinize the white plaintiff’s proffered
evidence far more skeptically, wondering whether the comparison set of workers was actually better treated or was truly comparable. By the same logic, two different triers of fact in the present day will evaluate a given claim of discrimination by reference to their own background assumptions about the societal pattern of discrimination against the relevant group. The legitimacy of these inferences is supported by naturalizing epistemology and Bayesian probability theory, the two basic themes within contemporary evidence scholarship.

Discrimination doctrine cannot be agnostic about the societal pattern of discrimination. It can only determine how and by whom that pattern is assessed. Those who believe that discrimination is not common will—and rationally should—impose stringent standards of proof, while those who believe that discrimination is common will—and rationally should—impose more relaxed standards. The use of these background assumptions is not bias or activism. It is a logically unavoidable part of the process of evaluating evidence.

For many years, assumptions about the societal pattern of discrimination entered the legal process through one of two backdoors. First, to the extent that a finding of discrimination was left to the trier of fact, background assumptions were merely an unstated factor determining the trier’s conclusions. Second, the Supreme Court also formulated burden of proof rules that took some discretion away from the trier of fact. These presumptions were not explicitly based on background assumptions about discrimination, but language in the opinions indicated some awareness by the Justices that background assumptions were relevant.¹

¹ See Part VII. A.1,
In the last fifteen years, though, the Supreme Court has retreated from formulating burdens of proof to govern findings of fact. It has thus assigned almost wholly to lower courts and triers of fact the task of determining whether discrimination has occurred. Since there is nothing close to a social consensus on the societal pattern of discrimination, the outcomes of cases are dismayingly inconsistent.

In recent years, a third avenue has developed for introducing background assumptions. Expert testimony about societal discrimination patterns has been presented in a small set of cases, class actions that challenge the use of subjective practices to evaluate employees and job applicants. Such testimony is the purest type of what is known as social framework analysis and has played a key role in the landmark Ninth Circuit case of *Dukes v. Wal-Mart*, the largest class action in U.S. history. The use of pure social framework evidence has been challenged as irrelevant and prejudicial. In light of the unavoidable role that background assumptions play, these objections are without force. Indeed, social framework evidence is relevant not only in cases challenging subjective practices but in every case in which the factual existence of differential treatment is in dispute.

---

2 See Part VII.A.2.

3 I will call social framework analysis “pure” when it includes only expert views on general social science research of the societal rate of discrimination and the conditions that facilitate discrimination. Another type of social framework analysis, “applied,” encompasses the inferences by experts from general research to the case at issue. For a discussion of this terminology and related issues, see Part I.


5 See cases cited note 37.
If background assumptions must be made—and they must be—they should be made openly and with appropriate consideration of relevant evidence. Social framework evidence promises to go some way to replacing the personal views of judges and juries with a more objective perspective. On the other hand, considerations of both consistency and judicial economy suggest the undesirability of turning each and every discrimination trial into a forum for examining the pattern of discrimination in the United States today.

Part II discusses the current case law on the role of the evidence of the societal rate of discrimination. Part III explains why the use of background assumptions is logically unavoidable. Part IV describes the cognitive psychology of background assumptions, or how real people make use of background assumptions when they make decisions. Part V explains why social framework evidence of societal patterns of discrimination meets the basic standards for admissibility. Part VI addresses the institutional concerns surrounding the use of social framework evidence in discrimination trials. Part VII proposes an agenda for incorporating social frameworks into discrimination doctrine, as well as several specific systems of presumptions that might be useful given various different conclusions courts might reach about the societal pattern of discrimination.

II. THE SOCIETAL PATTERN OF DISCRIMINATION IN CASE LAW

The societal pattern of discrimination is widely understood to be relevant to setting broad policy towards discrimination.6 The views of Supreme Court nominees on the

6 Views on factual matters such as the prevalence of discrimination may not be based entirely on factual grounds, but may be colored by the individual’s normative perspective. This issue is explored in the important literature on “cultural cognition.” See, e.g., Dan M. Kahan & Donald Braman, Cultural Cognition and Public Policy, 24 YALE L. & POL'y REV. 149 (2006); Dan M. Kahan, David A. Hoffman & Donald Braman, Whose Eyes are You Going to Believe - Scott v. Harris and the Perils of Cognitive Illiberalism, 122 HARV. L. REV. 837 (2008); Secunda, Paul M.,
nature and prevalence of societal discrimination are therefore subject to detailed and sometimes brutal inquiry.\textsuperscript{7} Presumptively conservative justices are expected to recognize that discrimination remains a continuing concern,\textsuperscript{8} while presumptively liberal judges are pressed to acknowledge that progress has been made in reducing discrimination.\textsuperscript{9}

---


\textsuperscript{7} The joint role of values and background assumptions in the assessment of particular case facts was expressed by Senator Durbin during Justice Sonia Sotomayor’s confirmation hearings:

\begin{quote}
\textbf{SENATOR DURBIN:} Study after study has shown that men and women on the bench sometimes ruled differently against discrimination cases. This doesn't mean their rulings are based on personal bias. It simply acknowledges that Americans see the world through a prism of various experiences and perspectives.
\end{quote}

\begin{quote}
Sen. Patrick J. Leahy Holds a Senate Judiciary Committee Meeting on the Nomination of Judge Sonia Sotomayor to be an Associate Justice of the Supreme Court of the United States, \textit{Congressional Quarterly}, July 28, 2009.
\end{quote}

\textsuperscript{8} During Chief Justice Roberts’ confirmation hearings, Senator Edward Kennedy stated:
SENATOR KENNEDY: . . . In the only documents that have been made available to us, it appears that you did not fully appreciate the problem of discrimination in our society. . . [Y]ou simply did not grasp the seriousness of the impact of discrimination on our country as a whole.

Later in the hearings, Senator Biden asked:

SENATOR BIDEN: OK. Judge, is gender, as you've written in a memo, a perceived problem or is it a real problem? . . .

JUDGE ROBERTS: . . . Of course gender discrimination is a serious problem.

U.S. Senator Arlen Specter (R-PA) Holds Hearings on Roberts Nomination, CONGRESSIONAL QUARTERLY, September 13, 2005.

During her confirmation hearings, Justice Sotomayor was questioned by Senator Kohl:

SENATOR KOHL: Judge, first, I'd like to discuss the issue of affirmative action. We can all agree that it is good for our society when employers, schools, and government institutions encourage diversity. On the other hand, the consideration of ethnicity or gender should not trump qualifications or turn into a rigid quota system. . .

JUDGE SOTOMAYOR: It is firmly my hope . . . that in 25 years, race in our society won't be needed to be considered in any situation. That's the hope. And we've taken such great strides in our society to achieve that hope.

Sen. Patrick J. Leahy Holds a Senate Judiciary Committee Meeting on the Nomination of Judge Sonia Sotomayor to be an Associate Justice of the Supreme Court of the United States, CONGRESSIONAL QUARTERLY, July 15, 2009.
Title VII prohibits employers from “discriminat[ing] against any individual . . . because of . . . race, color, religion, sex, or national origin.”  

For most purposes, the term “discriminate” means simply to treat differently, although Title VII raises a few definitional questions about what constitutes discrimination. Sometimes these definitional issues require assumptions about broad empirical questions other than the prevailing pattern of societal discrimination. Other definitional issues require courts to evaluate the costs and benefits of measures to erase past discrimination or otherwise accommodate the special needs of one group. Policy considerations, and therefore views on the pattern of societal discrimination, clearly come into play in determining the proper scope of accommodation under Title VII.

The typical discrimination case, however, does not raise these subtle doctrinal questions but depends merely on the factual question of whether the applicant or employee was subject to differential treatment “because of” membership in a protected class. In the overwhelming majority of Title VII discrimination cases, the plaintiff’s evidence does not unambiguously establish differential treatment based on protected class

---


membership. Instead, the plaintiff produces comparisons to the employer’s treatment of other employees; evidence that the employer’s stated reason was false; statements evincing improper attitudes; and other evidence that may be strong but that does not completely foreclose the possibility of non-discrimination.

Evidence about the societal pattern of discrimination has played no role in the overwhelming majority of employment discrimination cases, but in recent years, some plaintiffs have introduced evidence intended to establish that discrimination remains common and that it is likely to occur in circumstances like those found in the defendant firm. Such evidence is generally called social framework analysis, a term first introduced in a foundational article by John Monahan and Laurens Walker. In discrimination cases, social framework analysis is most often introduced in class actions, and the circumstances identified as problematic usually involve personnel systems with a subjective element. Although initially social framework evidence was introduced by plaintiffs, defendants have increasingly responded with their own expert witnesses.

---

14 The problems created by proof issues are discussed in Part VII.


A prominent and typical example of social framework analysis is the expert report of William Bielby in *Dukes v. Wal-Mart*.\textsuperscript{18} Bielby is both a leading scholar on


Outside of class challenges of subjective practices, social framework analysis has played a relatively small role in discrimination cases. Social framework evidence has been used to question the use of subjective procedures in most of the few individual cases in which such evidence has been admitted, including the landmark case *Price Waterhouse v. Hopkins*, Price Waterhouse v. Hopkins, 490 U.S. 228 (1989). See also Tuli v. Brigham & Women's Hospital, *supra* note\textsuperscript{17}. For unsuccessful efforts to use social framework analysis in individual cases, see Chadwick v. Wellpoint, *supra* note\textsuperscript{17}; Ray v. Miller Meester Advertising, *supra* note\textsuperscript{17}.

\textsuperscript{17} Ellis v. Costco Wholesale Corp., *supra* note\textsuperscript{17} (Dr. Margaret Stockdale and Dr. Casey Mulligan on women's lack of interest in relevant job; Dr. Frank Landy on non-discriminatory nature of practices); EEOC v. Morgan Stanley, *supra* note\textsuperscript{17} (Dr. June O'Neill on lack of interest. Dr. Barbara Gutek on employment practices, and Dr. Ira Kay on employment practices); Puffer v. Allstate Ins. Co., *supra* note\textsuperscript{17} (Dr. Christopher Winship on stereotyping).

discrimination\(^{19}\) and a leading expert for plaintiffs.\(^{20}\) Because of the case’s unprecedented size, both it and Bielby’s testimony have attracted a great deal of commentary.\(^{21}\) Bielby’s report described social science research on the overall societal pattern of discrimination and concluded that this research indicated that bias is common in contemporary society.

In the employment context, career barriers resulting from gender stereotypes and gender bias are likely to be consequential for women working in a traditionally male domain, such as the middle to upper managerial and professional ranks of large corporations, engineering divisions of firms, in the military, and in historically male-dominated industries such as skilled crafts and construction trades.\(^{22}\)

Bielby also described the social science research on the organizational factors that fail to check discrimination:


\(^{20}\) To honor his ubiquity, I propose naming motions to exclude such evidence “Bielby motions.”


\(^{22}\) *Dukes v. Wal-Mart Stores*, Expert Witness Report, Dr. William Bielby, *supra* note___ at text at note 44.
A large body of social science research demonstrates that stereotypes are especially likely to influence personnel decisions when they are based on informal, arbitrary, and subjective factors. In such settings, stereotypes can bias assessments of a woman's qualifications, contributions, and advancement potential, because perceptions are shaped by stereotypical beliefs about women generally, not by the actual skills and accomplishments of the person as an individual.23

Bielby’s testimony on the societal pattern of discrimination and on the institutional features that check discrimination drew on social science findings that are general in nature.24 Though intended to aid the jury in interpreting adjudicative facts, it did not purport to be based on the particulars of employment at Wal-Mart.25 Other expert witnesses have likewise testified about the general societal pattern of discrimination, apart from its application to case-specific facts.26

Bielby also testified about the applicability of general social science findings to Wal-Mart’s personnel procedures. His Summary of Findings stated:

23 Id. at text at note 45-46.
24 Dukes v. Wal-Mart Stores, Expert Witness Report, Dr. William Bielby, supra note___.
Subjective and discretionary features of the company's personnel policy and practice make decisions about compensation and promotion vulnerable to gender bias . . . I have concluded that there are significant deficiencies in the company's policies and practices for identifying and eliminating barriers to equal employment opportunity at Wal-Mart.27

Like Bielby’s report in Dukes, expert reports in other cases have often used general social science evidence as a basis for drawing conclusions about the practices of the defendant employer and sometimes even about whether discrimination has occurred in a the defendant firm.28 Such inferences were not part of social framework analysis as originally conceived by Monahan and Walker, which included only expert testimony on general social science research.29 That general research was intended to help the jury interpret case-specific evidence in order to reach factual conclusions.30 Monahan and Walker regard relatively informal interpretations of case-specific facts as inappropriate in expert testimony, suggesting that experts should only offer case-specific opinions reached through the rigorous application of social science methods to the particular facts of the case.31 Subsequently the term “social framework” has been widely used by Bielby and others to encompass both general social science research and the relatively informal

27 Dukes v. Wal-Mart Stores, Expert Witness Report, Dr. William Bielby, supra note___.


29 Contextual Evidence of Gender Discrimination, supra note___, at 1742-48, discussing Social Frameworks

30 Contextual Evidence of Gender Discrimination, supra note___.

application of that research to the case at issue.\textsuperscript{32} This broader usage has become widespread enough that I adopt it here, though this adoption does not itself imply any conclusions about which types of social framework evidence should be admissible. I use the expression “pure social framework” to refer to the original Monahan-Walker conception, and the term “applied social framework” to refer to applications of social frameworks to case facts that are made informally, without the use of rigorous research methods.

Some defendants have vigorously challenged the admissibility of pure social framework analysis,\textsuperscript{33} and these efforts have met with some success in court. In Parts III though V, I argue that pure social framework evidence meets admissibility requirements in an even wider range of cases than those in which it has so far been used.

III. THE LOGICALLY UNAVOIDABLE USE OF BACKGROUND ASSUMPTIONS

The threshold requirement of evidence law is relevance.\textsuperscript{34} Pure social framework evidence of the social pattern of discrimination has been challenged as irrelevant to the


\textsuperscript{33} Although both pure and applied social framework analyses raise some important issues, the focus of this article is on pure social framework evidence of the societal pattern of discrimination.

\textsuperscript{34} \textit{Federal Rules of Evidence} 401, \textit{Definition of "Relevant Evidence"}; \textit{Federal Rules of Evidence} 402, \textit{Relevant Evidence Generally Admissible; Irrelevant Evidence Inadmissible}. 
resolution of employment disputes that turn only on questions of fact. Only case specific facts, it might seem, are needed to determine whether differential treatment has occurred in a specific dispute. Yet specific disputes are always evaluated against background assumptions about the societal pattern of discrimination. This Part will provide first an intuitive and then a more rigorous explanation for the role of background assumptions in case-specific determinations.

A. Some Intuition

The relevance of the societal pattern of discrimination to case-specific factual questions is easiest to see in subjective practice cases. Even if subjective employment procedures frequently result in unreasonable and unfair outcomes, that unfairness is not discriminatory if it does not fall disproportionately on a protected class. Only if supervisors have a systematic bias against a particular group will unconstrained discretion tend to produce outcomes that discriminate against that group.

The relevance of societal discrimination patterns is less obvious in a typical discrimination suit that does not directly challenge subjective practices. In such cases, plaintiffs try to prove their case through various types of evidence including comparisons of their treatment to that of employees outside the protected class; evidence that the employer’s stated reasons for its action were false; and employer statements indicating stereotyping or discriminatory animus.

Before examining how background assumptions function in discrimination cases, consider a simpler situation. A patient who has either X, an extremely common disease or Y, an extremely rare one, will always have a positive result on a certain medical blood test. If a patient’s test is positive, which is more likely, that the patient has X or Y? Clearly the answer is X. Our knowledge that X is more likely than Y properly influences our interpretation of the test results.

The intuition behind discrimination cases is more subtle but the role that background assumptions play in assessing evidence is obvious in extreme cases. Suppose you know for sure that all employers are discriminating. You will require no evidence at all to convince you that differential treatment is occurring in a particular case. Conversely, suppose that you know for sure that differential treatment never occurs. No evidence can or should convince you that it is occurring in a particular instance. Now suppose you know that differential treatment is extremely common, but not universal. Very little evidence will be needed to convince you of differential treatment in a specific case. Again the converse holds. If differential treatment is known to be very rare, then only extremely strong evidence will convince you that differential treatment has occurred in a specific case. Clearly your background assumptions about the societal pattern of discrimination influence your interpretation of other evidence, as they should. Thus, evidence of the societal pattern itself is relevant and, absent other considerations, should be admissible.

We can further develop our intuitions about the relevance of background assumptions by imagining two cases, one involving a high probability of discrimination and the other a low probability. Consider a white-owned factory in the American South in 1965. In the high discrimination scenario, a black employee charges that the employer has discriminated. In the low discrimination scenario, a white employee in the same factory alleges reverse discrimination.
Now consider several classes of evidence that are commonly introduced by plaintiffs to prove discrimination. One important kind of evidence is comparative: the plaintiff compares his or her treatment with that of similarly situated employees outside the protected class. Several types of comparisons can be made. In promotion and hiring cases, the plaintiff’s qualifications are often compared with those of someone who was hired or promoted instead. In a discharge case, the plaintiff may try to show that individuals outside the protected class who displayed similar deficiencies were better treated.

A trier of fact evaluates these comparisons and the defendant’s responses to them against the background of some view about how common discrimination is. In the Southern factory example, suppose that a black plaintiff who was discharged for theft demonstrates that a white employee was not discharged for stealing the same amount under identical circumstances. Most of us would regard such evidence as highly probative of discrimination. We might greet with derisive laughter an employer claim that the black employee had been discharged because he was disliked by his white supervisor, but that this animosity was not racial in nature. Alternatively, suppose that the employer introduced evidence that the white employee not discharged for an identical theft had a marginally better performance record. The employer would have difficulty convincing us that the performance differences were sufficient to justify differential treatment.

On the other hand, we would be far more skeptical of a reverse discrimination claim by a white plaintiff in the same factory who was discharged even though a black employee who had committed an identical theft was not. We would probably regard the

36 For an example of such a defense, albeit under different circumstances, see St. Mary's Honor Center v. Hicks, 509 U.S. 502 (1993).
employer’s evidence of even marginally weaker job performance by the plaintiff as material and explanatory of the difference in treatment. We would be inclined to credit an argument by the employer that the discharge was motivated in part by a supervisor’s non-racial dislike of the plaintiff. Our different reactions to the two scenarios are perfectly rational uses of our background knowledge of the relative likelihood of discrimination against blacks and whites in the South in 1965.

Another important strategy used by discrimination plaintiffs is to introduce evidence that the defendant’s explanation for its decision is false.\(^\text{37}\) For example, suppose that an employer claims that the plaintiff was discharged for theft, but produces no evidence that any theft had been alleged during the plaintiff’s employment. Plaintiff calls other employees who testify that they had heard no suggestion of wrongdoing by the plaintiff. With a black plaintiff, evidence that the employer’s explanation was false would be powerful evidence of discrimination. A Southern employer in 1965 who seemed to have manufactured a serious charge against a black employee would have enormous difficulty convincing us that no discrimination had occurred. We would not be so readily persuaded by a white plaintiff, even with similar evidence of an apparently false explanation. We would not leap to the conclusion that the employer had lied, but would wonder if the employer had gotten mixed up. Perhaps the discharge occurred for a different legitimate reason, but the employer confused the plaintiff with another employee. Perhaps the theft had occurred, and the employer had felt no need to document the reason for the discharge in a time and place where the at-will doctrine was strong. We would search hard for such explanations and it would be extraordinarily difficult to persuade us that this discharge, even if unfair, was racially motivated.

\(^{37}\) This is a classic “pretext case”, discussed at length in Part VII.A.
A third type of evidence is direct evidence of racial animosity. In the 1965 South, we would tend to believe a witness produced by our hypothetical black plaintiff who testified that the defendant had made racist remarks towards blacks. However, if the white plaintiff produced a witness who testified that the employer expressed animus against whites, we would suspect that the witness had ulterior motives or had somehow misunderstood the remarks in question.

These hypothetical situations illustrate the importance of background assumptions. However, they do not suggest the need for evidence bearing on those assumptions, since the societal pattern of discrimination in these circumstances is beyond reasonable dispute. Consider now two triers of fact in the present day confronted with a discrimination claim. Wide disagreement exists about the current pattern of societal discrimination, and two jurors evaluating a single case may begin with very different background assumptions. Since those assumptions play a critical role in evaluating case specific facts, any effort to force jurors to focus on case specific facts alone will fail, and will simply hide an important element of their reasoning.

B. A Formal View of Background Assumptions

The role of background assumptions can be analyzed formally using Bayes’ Theorem.\textsuperscript{38} For several decades, Bayesian probability theory has played an important though controversial role in evidence scholarship.\textsuperscript{39} Bayes’ Law provides a logical


\textsuperscript{39} Important early expressions of skepticism include Laurence H. Tribe, \textit{Trial by Mathematics: Precision and Ritual in the Legal Process}, 84 HARV. L. REV. 1359 (1970); Lea
constraint, valid in all widely used definitions of probability, on the relation between background assumptions, evidence and conclusions. This Part will provide an overview of the application of Bayes’ Law to the analysis of evidence in discrimination cases.  

To see the role of background assumptions formally, we will call the legal hypothesis to be proven \( h \). We can think of \( h \) not as the ultimate question of liability but as merely a specific fact at issue, such as whether the defendant treated the plaintiff differently from comparable employees outside the protected class. The probability of the hypothesis being true in any randomly picked case may be called either the unconditional probability of \( h \) or the prior probability of \( h \).\(^{41}\) The prior probability of \( h \) divided by the prior probability of not \( h \) (that is, of \( h \) not being true) is called the odds or odds ratio of \( h \). Each of us forms our odds ratio of \( h \) based on various relevant background assumptions and any new evidence we are presented with.

Suppose that there is some type of case-specific evidence \( e \) that tends to support an inference of hypothesis \( h \). Either by statistical method or intuition, we have some sense of the probability that \( e \) would be observed if the hypothesis \( h \) were true (the probability of \( e \) given \( h \)). Similarly, we have some sense of the probability that \( e \) would be observed if \( h \) were not true (the probability of \( e \) given not \( h \)). Dividing the probability of \( e \) given \( h \) by the probability of \( e \) given not \( h \) yields the likelihood ratio, or the odds that \( e \) is true given \( h \). For example, suppose that \( h \) is the hypothesis of discrimination and that \( e \) consists of

---

\(^{40}\) A formal analysis of Bayes’ Law is given in Part IX.

\(^{41}\) In special cases, the prior probability can be summarized by the frequency with which \( h \) is true in a given population, and this frequency is known as the base rate of \( h \).
evidence that the plaintiff’s supervisor had stated openly that he did not like working with members of the plaintiff’s class. The likelihood ratio for this evidence would be high: the probability of this evidence given discrimination is much higher than the probability given non-discrimination.

In a legal proceeding, the likelihood ratio is not an end in itself but is a means to establish the odds that the hypothesis $h$ is true given evidence $e$, known as the posterior odds of $h$ given $e$. The posterior odds are the probability of $h$ given $e$ divided by the probability of $h$ given $\neg e$. The posterior odds and the likelihood ratio are two subtly but critically different numbers, though they are often confused.\(^{42}\) To obtain posterior odds, a likelihood ratio must be evaluated in light of the prior odds that $h$ is true without knowing $e$. Bayes’ Law shows a simple relationship between the posterior odds (the odds that $h$ is true given $e$), the likelihood ratio (the odds that $e$ is true given $h$), and the prior odds (the odds that $h$ is true without knowing $e$):

\[
\text{Posterior odds} = \text{Likelihood ratio} \times \text{Prior odds}
\]

Bayes’ Law shows that assumptions about prior odds are unavoidable in any inference from a likelihood ratio to posterior odds. Suppose that we were to decide that our estimate of the prior odds was unreliable and thus should be ignored. We disregard the prior odds and assume that the posterior odds are equal to the likelihood ratio. Bayes’ Law implies that this would amount to an assumption that the prior odds ratio is 1.\(^{43}\) A prior odds ratio of 1 in turn implies that discrimination and nondiscrimination are equally

---


\(^{43}\) Posterior odds = likelihood ratio * 1
probable.\textsuperscript{44} Our efforts to disregard the prior odds of discrimination have failed. We have simply substituted an assumption of equal probability for an empirical estimate.

\textit{C. Class actions}

In discrimination class actions, statistics are central to proof of discrimination.\textsuperscript{45} In a class discrimination case, \( h \) is the hypothesis that members of protected class were treated differently from other employees, while \( e \) is an observed discrepancy between the proportion of some protected group in the employer’s workforce and in the relevant labor market.\textsuperscript{46} Of course, such disparities can be innocent. A trier of fact would like to know the odds that the disparity in question occurred because of differential treatment.

The likelihood ratio is relatively easy to estimate in these cases. Given the number of employees in the firm, the proportion of the protected group in the qualified population,

\textsuperscript{44} Posterior odds = likelihood ratio \(*\frac{5}{5}\)

\textsuperscript{45} The observation that explicitly statistical analysis of discrimination presumes prior probabilities has been made in D. Kaye, \textit{Statistical Evidence of Discrimination}, 77 J. AMERICAN STATISTICAL ASSOCIATION 773–783 (1982); Strangely Persistent Transposition Fallacy, \textit{supra} note___. These writers suggested that the introduction of evidence bearing on background assumptions was not feasible, a view I too might have shared in the past. However, I will argue that both social science research and its use in the courts have progressed to a point where evidence of background assumptions can perform a useful role. \textit{See infra} Part __. See generally David L. Faigman, \textit{The Tipping Point in the Law's Use of Science: The Epidemic of Scientific Sophistication that Began with DNA Profiling and Toxic Torts}, 67 BROOK. L. REV. 111 (2001).

\textsuperscript{46} Where a job requires specific skills, the qualified population consists of workers with those skills. Few jobs are filled by nation-wide search, and the relevant labor market is also limited by some measure of proximity to the employer. Hazelwood School District v. United States, 433 U.S. 299 (1977).
and the proportion of the protected group in the firm, a statistician can estimate the relative likelihood that the discrepancy resulted from discrimination or non-discrimination.\textsuperscript{47}

The current legal standard requires plaintiff to show that the chance that nondiscriminatory hiring produced the observed discrepancy is no greater than .05, which is approximately equal to a requiring a likelihood ratio of 20 or higher.\textsuperscript{48} If discrimination and non-discrimination are equally likely, a likelihood ratio of 20 implies a 95.2% probability of discrimination. A probability of 95.2% certainly seems sufficient (and perhaps more than sufficient) to meet a plaintiff’s initial burden, but depends on the assumption that discrimination and non-discrimination are equally likely. If only 2% of all firms discriminate, a likelihood ratio of 20 implies that the probability of discrimination is about 29%. This is an extremely light threshold for plaintiffs. Conversely, if far more than half of all firms discriminate, a 95.2% chance of discrimination will be implied by a likelihood ratio of much less than 20, and requiring a likelihood ratio of 20 may impose an excessively high burden on plaintiffs.

\textit{D. Informal Evidence: Individual Disparate Treatment}

The applicability of Bayes’ law is most obvious for evidence that is explicitly statistical. Formal statistical modeling, however, only makes explicit a mode of inference that is implicit in many evidentiary settings. Prior odds also play a critical role in evaluating non-statistical evidence, the principal means of proof in individual disparate treatment cases and also of importance in class actions.

\begin{footnotesize}
\textsuperscript{47} Details of this calculation are provided in the Appendix.

\textsuperscript{48} More detail is provided in the Appendix.
\end{footnotesize}
Consider the earlier hypothetical in which an employee is discharged for stealing. The plaintiff demonstrates that another employee was not discharged for an identical theft, and the defendant establishes that the other employee has a marginally better performance record. These facts together are a set of evidence $e$. Each of us has a hunch about the relative likelihood of such a scenario with and without discrimination. Just as in the statistical case, we can approximate a value for the likelihood ratio. I personally suspect that the likelihood ratio is high. Employers, I would imagine, take theft more seriously than marginal performance differences. Small performance differences would not justify retaining an employee who committed a theft that was large enough to justify another employee’s discharge. Thus, I would expect that the majority of cases yielding this evidence would involve discrimination. Whether you share this intuition is not critical: all that matters is that you have some view of the relative likelihood of this scenario with and without discrimination.

Suppose that we estimate that retaining another employee who stole but had marginally better performance is 20 times more likely with discrimination than without it—in other words, the likelihood ratio is 20. Plaintiff thus appears to have a very powerful case. Yet we do not and should not evaluate this evidence in a vacuum. If the plaintiff in question is a black Southern factory worker in 1965, we would be inclined to conclude that discrimination had occurred, while if the plaintiff is a white worker in the same plant we would require more evidence, if indeed any evidence at all would convince us. The difference in our intuitions results from our different beliefs about the prior odds ratio. Even if we assume that discrimination played a role in only 10% of all discharges of black employees (an assumption generous to defendants) a likelihood ratio of 20 would mean that discrimination is more than twice as likely as non-discrimination. With a white plaintiff at that time and place, a 10% prior probability of race discrimination seems implausibly high. Even a prior probability of one-tenth of one
percent, (.1%) seems generous to plaintiffs, and with that background probability, the likelihood ratio of 20 implies that non-discrimination is fifty times more likely than discrimination.

Similar reasoning applies to other kinds of informal evidence, such as testimony that the defendant displayed discriminatory attitudes or that the defendant’s stated reason for discharge was factually false. All evidence, just like statistical data, must be combined with the prior odds of differential treatment in order to reach a conclusion about the likelihood of discrimination in a particular case. Any attempt to ignore prior odds simply amounts to an assumption that discrimination and non-discrimination are equally likely.

IV. THE COGNITIVE PSYCHOLOGY OF BACKGROUND ASSUMPTIONS

Evidence law is not written on a blank slate. The decision to permit or exclude evidence of societal discrimination patterns depends in part on how jurors are likely to use this evidence, which this Part will now examine.

A great deal of research has examined the extent to which people use evidence that bears on background assumptions in their reasoning processes and the extent to which their usage conforms to the Bayesian model. Two diverging lines of research have

49 The role of prior probabilities in pretext cases is noted in Linda Hamilton Krieger, Burdens of Equality: Burdens of Proof and Presumptions in Indian and American Civil Rights Law, 47 AM. J. COMP. L. 89, 120 (1999).

50 The examples given above are numerical and might seem open to the objection that none of us have quantified views on the probabilities in question. However, Bayes’ Law also holds for qualitative judgments about relative likelihood although the formal representation of these relationships is more complex, and the numerical example is a simpler way of showing the general argument. P. Suppes, Qualitative theory of subjective probability, in SUBJECTIVE PROBABILITY (1994).
reached apparently contradictory conclusions. In some contexts, people appear to underweight prior odds in comparison to the Bayesian norm.\footnote{This phenomenon was initially noted in P. E. Meehl & A. Rosen, \textit{Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores}, 52 \textit{Psychological Bulletin} 194–216 (1955); Daniel Kahneman & Amos Tversky, \textit{Subjective probability: A judgment of representativeness}, 3 \textit{Cognitive Psychology} 430–454 (1972). For an excellent survey of the early research finding base rate neglect, see M. Bar-Hillel, \textit{The base-rate fallacy in probability judgments}, 44 \textit{Acta Psychologica} 211–233 (1980).} In other contexts, evidence of so-called confirmation bias indicates that people place great weight on prior odds and update those odds far less than Bayesian norms would seem to dictate.\footnote{For excellent surveys, see R. S. Nickerson, \textit{Confirmation bias: A ubiquitous phenomenon in many guises}, 2 \textit{Review of General Psychology} 175–220 (1998); Joshua Klayman, \textit{Varieties of Confirmation Bias}, Volume 32 \textit{Psychology of Learning and Motivation} 385-418 (1995).} These findings raise several questions. Why do results vary so greatly between studies? Are apparent departures from Bayesian norms evidence of irrational decision-making, or are experimenters failing to note some type of legitimate inference that subjects are using?\footnote{A turning point in the study of base rate psychology was a highly influential paper by Jonathan Koehler. Jonathan J. Koehler, \textit{The base rate fallacy reconsidered: Descriptive, methodological and normative challenges}, 19 \textit{Behavioral and Brain Sciences} 1–17 (1996). See also Jonathan J. Koehler, \textit{Issues for the next generation of base rate research}, 19 \textit{Behavioral and Brain Sciences} 41–53 (1996) (responding to commentators); Jonathan J. Koehler, \textit{A farewell to normative null hypothesis testing in base rate research}, 20 \textit{Behavioral and Brain Sciences} 780–782 (1997) (same). Prior to this paper, many researchers seemed overly inclined to interpret experimental results as an example of irrationality. Koehler had developed some of these ideas in earlier papers, Jonathan J. Koehler, \textit{Base rates and the "Illusion" Illusion}, 5 \textit{Psycoloquy} (1994); Jonathan J. Koehler, \textit{The base rate fallacy myth}, 4 \textit{Psycoloquy} (1993).}
Subsequent work has in part answered these questions by examining the use of prior odds with more attention to context, content, and mode of presentation.\textsuperscript{54} In some instances, people do truly underweight prior odds and depart at least somewhat from any plausible standard of rationality. This neglect of prior odds is typically found when subjects form prior odds based only on evidence within the experimental setting. Underweighting of prior odds has been found when subjects are presented with both legal and non-legal material.\textsuperscript{55} These deviations are reduced, sometimes significantly, when information about prior odds is presented in a cognitively familiar format. For example, most people have difficulty using information presented as a numerical probability but are far more Bayesian when provided with information in a frequency format.\textsuperscript{56} Similarly, most of the studies finding relatively high neglect of prior odds involve

\begin{itemize}
  \item \textsuperscript{54} The importance of more nuanced and ecologically valid studies was emphasized by Koehler. The base rate fallacy reconsidered, \textit{supra} note__.
  \item \textsuperscript{56} G. Gigerenzer et al., \textit{Cognitive illusions reconsidered}, 1 \textsc{Handbook of Experimental Economics Results} 1018–1034 (2008); U. Hoffrage et al., \textit{Representation facilitates reasoning: What natural frequencies are and what they are not}, 84 \textsc{Cognition} 343–352 (2002); Gerd Gigerenzer & Ulrich Hoffrage, \textit{The Role of Representation in Bayesian Reasoning: Correcting Common Misconceptions}, 30 \textsc{Behavioral and Brain Sciences} 264-267 (2007); Other factors besides frequency format may play a role. Aron K. Barbey & Steven A. Sloman, \textit{Base-Rate Respect: From Ecological Rationality to Dual Processes}, 30 \textsc{Behavioral and Brain Sciences} 241-254 (2007); L. Macchi, \textit{Pragmatically before ecologically valid tasks}, 20 \textsc{Behavioral and Brain Sciences} 778–779 (1997); V. Girotto & M. Gonzalez, \textit{How to elicet sound probabilistic reasoning: Beyond word problems}, 30 \textsc{Behavioral and Brain Sciences} 268–268 (2007).
\end{itemize}
abstract and unfamiliar problems. Subjects come closer to the Bayesian norms in solving problems that involve familiar objects or situations, especially humans.\textsuperscript{57}

In contrast to prior odds described within the experiment, prior odds derived from the subject’s pre-laboratory background assumptions are given heavy weight and are relatively resistant to new evidence.\textsuperscript{58} Unsurprisingly, this confirmation bias is especially powerful for background assumptions that are self-interested or that form an important component of the subject’s worldview.\textsuperscript{59} The term “bias” may be misleading in this setting: Sometimes confirmation bias is rational, while at other times it is not. No algorithm dictates how people must arrive at a likelihood ratio, and the prior odds may logically be one relevant factor.\textsuperscript{60} Unyielding refusal to consider new evidence surely

\textsuperscript{57} Lívia Markóczy & Jeffrey Goldberg, \textit{Women and Taxis and Dangerous Judgments: Content Sensitive Use of Base-Rate Information}, 19 \textsc{Managerial and Decision Economics} 481-493 (1998).


\textsuperscript{59} E. L Uhlmann, V. L Brescoll & D. Pizarro, \textit{The motivated use and neglect of base rates}, 30 \textsc{Behavioral and Brain Sciences} 284–285 (2007); Confirmation bias, \textit{supra} note \textsuperscript{58}.

\textsuperscript{60} Jonathan J. Koehler, \textit{The influence of prior beliefs on scientific judgments of evidence quality}, 56 \textsc{Organizational Behavior and Human Decision Processes} 28–28 (1993); A. Gerber & D. Green, Misperceptions about perceptual bias, 2 \textsc{Annual Review of Political Science} 189 (1999).
sometimes leads to disasters. On the other hand, pre-existing prior odds may be the product of careful synthesis of a large volume of information, and any single new piece of information may not warrant significant updating of those prior odds. At the present time there is no general theory for distinguishing when confirmation bias is reasonable and when it is not.

Subjects give much higher weight to prior odds, however learned, when a plausible causal mechanism connects the information from which those odds are formed to the facts at issue. Studies of the behavior of judges and juries suggest that causal hypotheses play an important role in their use of background evidence. For example, suppose that blood found on a crime scene is type A, the same type as the defendant's. Evidently because there is no causal relation between blood type and criminal propensity, the prior odds of the observed blood type in the population usually receives little weight unless the blood type is rare, such as AB negative. Conversely, evidence of prior bad acts has some causal explanatory power and thus tends to receive substantial and perhaps excessive weight. Early observers tended to regard this use of causal explanations as irrational, while more recent commentary typically regards this as a rational approach.

\[61\] Jonathan J. Koehler, When Do Courts Think Base Rates Statistics Are Relevant?, 42 JURIMETRICS 373 (2001); N. Pennington & R. Hastie, A Cognitive Theory of Juror Decision Making: The Story Model, 13 CARDOZO L. REV. 519 (1991); Bayes' theorem in the trial process, supra note ___. Courts also seem to find evidence of prior odds especially compelling when no alternative cause is offered. When Do Courts Think Base Rates Statistics Are Relevant?, supra note ___.


A juror’s prior odds regarding the societal pattern of discrimination meets all of the conditions for high use of background assumptions: they are based on pre-courtroom experience; they concern familiar human behavior; and they are often based on strongly held views. It thus seems likely that these background assumptions play an important role in reasoning about discrimination. However, the typical person’s reasoning probably cannot be reduced to a simple application of Bayes’ Law and will consider factors such as the plausibility of non-discriminatory explanations for employer conduct.

V. LEGAL EVIDENCE BEARING ON BACKGROUND ASSUMPTIONS

The two threshold questions in evidence law are whether evidence is relevant and whether it is substantially more prejudicial than probative. This Part examines whether evidence bearing on background assumptions, or background evidence, meets traditional admissibility requirements. Neither empirical studies of how people use background
assumptions nor Bayesian inference rules are in themselves sufficient to answer these questions, though both influence the answers. This Part will examine the canonical scholarly examples of background evidence; some philosophical frameworks for evidence law; and the application of those frameworks to analysis of the relevance and prejudicial effect of background evidence.

A. The Canonical Comparisons

The admissibility of evidence about background assumptions has been the subject of much controversy in legal scholarship. However, the canonical background evidence differs in an important way from evidence about the societal pattern of discrimination.

The well-known Blue Bus hypothetical is frequently used to illustrate general problems with background evidence. Suppose that the plaintiff is hit by a bus that flees

66 See, e.g. R. C. Park et al., Bayes Wars Redivivus—An Exchange, 8 INTERNATIONAL COMMENTARY ON EVIDENCE 1 (2010).


68 In the similar paradox of the gatecrasher, we imagine that 499 people have paid for admission to a sporting event, and 1000 people are observed in the stands, so that 501 must have snuck in. If there is no other evidence, then there is a just over 50% chance that each person in the stands snuck in. If we admit this statistical evidence and we interpret the civil standard of allowing recovery as requiring a barely over 50% chance of wrongdoing, then the stadium owner can recover against each person in the stands. L. J. COHEN, THE PROBABLE AND THE PROVABLE 75 (1977). Another hypothetical concerns twenty-six prisoners, whose names are all known, in a prison yard. One is killed by twenty-four of the others, while the remaining prisoner hides. A witness viewed these events at a distance and cannot identify the non-participant. May a
the scene of the accident. An eyewitness sees that the bus is blue but can provide no other identifying information. The Blue Bus Company, whose buses are appropriately painted, owns 80% of the blue buses in town. Can the plaintiff recover based on that evidence alone?

An important difference must be noted between the evidence of the societal pattern of discrimination and case-specific background evidence as in the Blue Bus hypothetical. Where background evidence is case-specific, juries can effectively be led by evidentiary exclusion to use assumptions other than the descriptively accurate ones, and some policy might be advanced by those alternative assumptions. Consider the general rule excluding past convictions and other character evidence. In the absence of evidence of past convictions, juries will assume that the defendant had the same propensity to act in a specific way as a typical person with the characteristics observed in admissible evidence. In the Blue Bus cases, we can effectively exclude evidence of background facts if we choose to do so, and this exclusion will alter the background assumptions that juries make in a way desired by those opposed to the use of background evidence. If evidence of the proportion of blue buses owned by defendant is excluded, juries will probably assume that the ownership of blue buses is widely dispersed and assign little weight to the fact that a witness saw a blue bus. The evidentiary exclusion seems to achieve whatever goal is thought to be served by focusing on case-specific evidence. In contrast, juries enter the courtroom with preexisting views about the societal pattern of discrimination and cannot be forced by the fiat of evidentiary exclusion to turn off their views about that societal pattern. Excluding evidence of societal patterns of

prosecutor using only this naked statistical evidence charge all twenty-five with murder? The criminal setting of these examples raises difficult issues not relevant to the present civil setting.
discrimination will not eliminate the use of background assumptions, but will merely substitute the jury’s pre-existing views for views supported by evidence.

The exclusion of evidence bearing on the usual types of prior odds may succeed in shaping jury assumptions in some desired fashion. In contrast, any effort to exclude evidence of the societal pattern of discrimination will simply leave pre-courtroom background assumptions unchanged. With this distinction in mind we can examine how general arguments for and against the admissibility of background evidence applies to evidence of the societal pattern of discrimination.

**B. Philosophical Foundations of Evidence Law**

Much contemporary evidence scholarship draws on epistemology, the study of how we arrive at true belief. An important theme in contemporary epistemology, summarized in the expression “naturalized epistemology,” is that evidence of how people actually think must play an important role in the study of the justification of knowledge. Extreme naturalism virtually equates epistemology with cognitive psychology and seems to abandon any inquiry into whether our beliefs are true. A less radical and more

---


70 Replacement naturalism, the original and most extreme version, was originally advanced by W.V.O Quine, W. V. O. Quine, *Naturalized Epistemology*, in *ONTOLOGICAL RELATIVITY AND OTHER ESSAYS* 69–90 (1969). It seems to preclude epistemology from examining the normative
widespread version of naturalism, normative naturalism, embraces a central role in epistemology for empirical science generally and cognitive science in particular while rejecting the view that nothing can be said about the truth or falsehood of our beliefs.

One of the most important and influential proponents of this approach is Alvin Goldman, whose work is of special interest to legal scholars because of the extensive attention he has devoted to social institutions generally and problems of legal evidence in particular. Epistemics, his suggested term for the revised enterprise of epistemology, seeks to regulate or guide our intellectual activities.\textsuperscript{71} This task requires a description of how we actually approach problems,\textsuperscript{72} an understanding of how these approaches may fail,\textsuperscript{73} and a theory of how these approaches might realistically be improved.\textsuperscript{74} Goldman distinguishes three ways of forming beliefs. Primary individual methods are basic or native cognitive processes such as perception, memory and reasoning.\textsuperscript{75} Secondary individual methods are learnable algorithms or procedures for arriving at beliefs.\textsuperscript{76} They problem of ascertaining whether our beliefs are true, or at least justified, in favor of the descriptive enterprise of examining why we believe what we believe. Jaegwon Kim, \textit{What Is "Naturalized Epistemology?"}, 2 PHILOSOPHICAL PERSPECTIVES 381-405 (1988). Few are ready to accept this result, and replacement naturalism has few advocates—indeed, Quine himself later qualified his views.


\textsuperscript{72} Epistemics, supra note __, at 511-17.

\textsuperscript{73} Epistemics, supra note __, at 517-20.

\textsuperscript{74} Epistemics, \textit{supra} note __, at 20.

\textsuperscript{75} Alvin I. Goldman, \textit{Foundations of social epistemics}, 73 SYNTHSE 109, 131 (1987); See generally \textit{ALVIN I. GOLDMAN, EPISTEMOLOGY AND COGNITION} (1986).

\textsuperscript{76} \textit{Foundations of social epistemics}, \textit{supra} note __, at 131.
include general topic-neutral techniques such as mathematics and logic as well as laboratory methods specific to particular scientific disciplines. Finally, social methods appeal, in whole or part, to the assertions, utterances, and beliefs of other people.\textsuperscript{77} The task of epistemics is to select from these three groups the practices that are most likely to lead to truth. No one set of practices is privileged. The best practice will vary depending on goals, circumstances and agents. The process of choosing practices is iterative and there is no guarantee that consensus among agents will be reached or that any consensus will be optimal. However, there are some forces that tend to produce convergence towards both consensus and truth.\textsuperscript{78}

From the perspective of normative naturalism, evidence law is primarily though not exclusively a process of searching for truth.\textsuperscript{79} The first task of evidence law is to identify the primary individual methods that triers of fact are likely to bring to the process of evaluating evidence.\textsuperscript{80} Given these methods, the second task is to design a set of social practices— the legal system—that are most likely to lead to veritistic results.\textsuperscript{81}


\textsuperscript{80} Epistemics, supra note ___(noting importance in all settings of beginning with primary processes); Alvin I. Goldman, \textit{Simple heuristics and legal evidence}, 2 \textit{LAW, PROBABILITY AND RISK} 215 (2003) (providing examination of implications of specific primary processes for evidence law).

Secondary algorithmic practices may lead us to endorse legal practices intended to override primary practices, and this effort is a form of epistemic paternalism. On the other hand, the law of evidence should not attempt wholly to replace the jury's primary approaches with algorithmic ones. From a practical point of view such wholesale replacement is not possible: the average juror simply cannot be forced to abandon all use of primary processes. From a theoretical point of view no set of algorithmic methods to date fully captures all considerations that properly belong in the evaluation of legal facts.

This three-part distinction among practices is useful for comparing the approaches of evidence theorists, even those not self-identified as normative naturalists. Evidence theorists differ in the extent to which they defer to primary and social practices in designing the legal system. Ordinary human decision makers play a much greater role in the legal system than in scientific inquiry. Some legal scholars argue that the centrality of lay reasoning places severe limits on the value of algorithmic metrics in designing evidentiary rules. The rules of evidence should be designed through examination of primary and social epistemic methods. They suggest that the most important primary method in legal settings is to compare the relative plausibility of the explanations offered

---


83 Epistemic Paternalism, supra note ___.

84 Allen and Leiter argue, with considerable support, that most modern evidence scholars are *de facto* normative naturalists whether or not the self-identify as such. Naturalized epistemology and the law of evidence, *supra* note ___.

85 Id.
by the plaintiff and defendant, a process sometimes called Inference to the Best Explanation.

Other scholars agree that legal rules must work in conjunction with actual human thought processes, but suggest that rule design may be helped by an inquiry into algorithmic methods, especially Bayesian ones, that may end up playing no direct role in the social practices on which we ultimately settle. All forms of Bayesianism require that rational agents update their prior odds at least in part by using Bayes’ Law, but by itself Bayes’ Law does not tell us how prior odds are to be formulated. Bayesianism refers to several classes of broader theories that attempt to answer this and related questions. It is not a unified set of views but a “theme that admits of many variations.”

86 Id.


90 Probability, logic and probability logic, supra note ___.
Orthodox Bayesianism provides a virtually exclusive role for secondary algorithmic processes. Although all Bayesian theories are “epistemic” and interpret probability as an estimate of the state of our knowledge, orthodox Bayesianism goes further, regarding prior odds solely as a measure of subjective personal belief, without inquiring into how that belief was formed. Orthodox Bayesians further require that rational agents form beliefs only by updating through Bayes’ Law. Evidence of how people actually reason is irrelevant.

Orthodox Bayesians are probably quite rare among evidence scholars. Less orthodox Bayesians may allow prior odds to be formed, at least in part, from objective

---

91 Bayesian probability theory contrasts with frequentist approaches, which define the probability of an event as its relative frequency after repeating a clearly defined process such as a coin flip a large number of times. To the frequentist, unique events cannot be assigned probabilities, and the statement “John is probably the person who killed Jane” is meaningless. Either he did or he didn’t. To both subjective and objective Bayesians, the statement summarizes our knowledge about whether John killed Jane. Bayes’ Law is valid within frequentist theory. However, a great deal of useful information cannot readily be expressed as a frequency, thus limiting the use of frequentist statistics. The increased popularity of Bayesian approaches is primarily motivated by a desire to make use of such non-frequency information. E. T Jaynes, *Prior probabilities*, 4 IEEE TRANSACTIONS ON SYSTEMS SCIENCE AND CYBERNETICS 227–241 (1968).


information;⁹⁴ may endorse updating rules in addition to Bayes’ Law;⁹⁵ and may evaluate hypotheses not only through updated probabilities but also by other criteria such as whether the hypothesis is simple, elegant, coherent, plausible relative to other hypotheses, and productive of new hypotheses.⁹⁶ These additional principles sometimes come from the observation of primary or social processes.⁹⁷ Thus, less orthodox Bayesians are in many respects like normative naturalizers who place significant weight on algorithmic methods.

C. Relevance

All evidence must meet a threshold requirement of relevance. The standard of relevance is veritistic, that is, the evidence must tend to advance the pursuit of truth. In the words of the Federal Rules of Evidence, evidence is relevant if it has “any tendency to make the existence of any fact that is of consequence to the determination of the action more probable or less probable than it would be without the evidence.” Defendants regularly challenge as irrelevant evidence of societal discrimination patterns, and some


⁹⁷ Knowledge in a social world, supra note ___.
courts have accepted their arguments, evidently on the grounds that general social data do not bear on specific cases.

The sharpest philosophical disputes in evidence scholarship concern whether juries should be explicitly instructed in the use of Bayes’ Law to incorporate evidence of prior odds. Scholars who emphasize algorithmic methods take evidence that subjects neglect certain types of background evidence as a sign of imperfect rationality in need of correction, and advocate that juries be instructed on the use of Bayes’ Law. Those who emphasize primary methods generally believe that apparent disregard of background evidence indicates a more subtle grasp of the relevant inference than Bayes’ Law alone. Other scholars take an intermediate position, regarding some apparent violations of Bayes’ Law as the result of the wisdom of juries and other the result of failure of primary approaches.

Even the most enthusiastic advocates of instructing the jury on Bayes’ Law have focused on highly quantified evidence such as DNA matching. Evidence of societal discrimination patterns is quantitative in only the loosest sense: it attempts to show that discrimination is common or not common. In theory jurors might make use of this information in a Bayesian or non-Bayesian way but the line between the two would be hard to discern. Instructing the jury in the use of Bayes’ Law would serve no purpose here, whatever its merit in other situations.

Background evidence about societal discrimination patterns and about case-specific facts present very different issues. Jurors typically have settled views on societal

98 A Bayesian Approach to Identification Evidence, supra note __.

99 Review, supra note __.

discrimination patterns before they enter the courtroom, and evidence of confirmation bias suggests these background assumptions will be given significant weight in interpreting the evidence. Evidence of societal patterns of discrimination thus bears on a factor that jurors already use in assessing evidence, and it is therefore relevant. Of course, societal patterns of discrimination may be inadmissible for other reasons, such as reliability. The reliability of expert testimony is obviously a question of enormous importance, but it is a question quite distinct from relevance.¹⁰¹

Once the use of Bayes’ Law is not an issue, the differences between various epistemic approaches to background evidence seem to diminish.¹⁰² Cognitive psychology overwhelmingly supports the view that people make use of background assumptions, regardless of whether they do so in a purely Bayesian fashion.¹⁰³ Thus, evidence regarding background facts seems to meet the relevance requirement of more or less all flavors of normative naturalism: from the naturalist perspective the suppression of background assumptions is psychologically futile while from the Bayesian perspective it is logically futile.

¹⁰¹ See Part V.E.

¹⁰² However, different epistemic approaches may lead to different conclusions about the relevance of evidence other than background facts. The evidentiary framework of discrimination law gives a central role to competing explanations for the defendant’s actions. Normative naturalists who stress the role of primary methods in the legal system should support this approach, since research strongly suggests that jurors evaluate evidence not only probabilistically but by comparing the relatively plausibility of two stories. Eclectic Bayesians may also endorse a role for relative plausibility in hypothesis evaluation while orthodox Bayesians will reject such a role.

¹⁰³ See Section IV.


D. Prejudice and Probative Value

Background assumptions play a role both in the cognitive processes of actual humans and in the Bayesian framework, and thus from most epistemic perspectives, evidence about the societal pattern of discrimination should meet the test of relevance. Still, relevant evidence may sometimes be inadmissible, especially when “its probative value is substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury.”

Evidence of societal patterns of discrimination has been excluded as prejudicial on the grounds that it impermissibly shifted the burden of proof to defendants by creating a presumption of discrimination. This use of the phrase “burden of proof” seems inconsistent with its legal meaning. In ordinary English, any powerful piece of evidence—a video seemingly showing the defendant committing the crime—might be said to shift the burden of proof. In the technical legal sense, though, the evidence does not shift but instead (arguably) meets the prosecution’s burden of proof. To say that a particular piece of evidence meets the burden of proof in an impermissible way simply raises the question of why the evidence is impermissible.

Several stronger objections to the use of background evidence have been made in the legal scholarship. These concerns of unfair prejudice fall into two groups: concerns that

\[104\] Federal Rule of Evidence 403

\[105\] See, e.g., EEOC v. Morgan Stanley, supra note___ (excluded in part, admitted in part); EEOC v. Wal-mart, supra note___ (excluded).
background evidence does not lead us closer to the truth and concerns that this evidence leads to valid inferences but is inconsistent with some other non-veritistic policy.\textsuperscript{106}

An important veritistic concern raised about background evidence is that imperfectly rational jurors will give that evidence excessive weight. Laboratory studies suggest that jurors are in fact likely to give the background assumptions with which they enter the courtroom great and perhaps excessive weight.\textsuperscript{107} Indeed, if this were not the case, jury selection would be a much simpler matter. But background evidence provided at trial will typically be under-weighted in comparison to Bayesian norms. At most, it is a partial antidote to the over-weighting of pre-trial background assumptions. To be sure, subjective juror background assumptions are not wholly without value, and social science evidence may be flawed. But any consideration of whether juror preconceptions are preferable to social framework evidence goes to the reliability of the evidence, not its relevance.\textsuperscript{108}

A non-veritistic policy objection to the use of background evidence is that the goal of individualized justice requires the use of case-specific evidence. The Blue Bus hypothetical is intended to illustrate this concern: evidence of the 80\% ownership rate, it is argued, is not sufficiently individualized. Whatever the normative merits of this argument in the Blue Bus context,\textsuperscript{109} it assumes that background assumptions can be

\begin{flushright}
\textsuperscript{106} Veritistic and policy concerns are frequently conflated. The distinction between them is explored in an illuminating article by Jonathan Koehler and Daniel Shaviro. Veridical Verdicts, \textit{supra} note\underline{___}. \textsuperscript{107} See Part IV. \textsuperscript{108} See Part V.E. \textsuperscript{109} Errors resulting from faulty individualized evidence, such as eyewitness testimony, bother me neither more nor less than errors from statistical evidence. Suppose that the Blue Bus company owned all the blue buses in town, but that the visibility conditions at the time of the accident were
\end{flushright}
feasibly excluded from the jury’s thought processes. While this assumption may generally be true in the Blue Bus hypothetical, it does not apply to evidence about societal discrimination patterns. People will inevitably refer to their experiential background assumptions about societal discrimination. The legal system can play a role in shaping what background assumptions are made but not whether jurors employ them.

Some misgivings about background evidence involve the highly quantified nature of certain types of background evidence, as in the Blue Bus case. Lawrence Tribe, for not perfect. Exhaustive studies of eyewitness evidence suggest an 80% likelihood of accurate color identification under the circumstances present. The eyewitness evidence is individualized, but this evidence and the original hypothetical statistical evidence have equal potential to produce erroneous findings of liability. As a practical matter, no system of civil justice can require certainty or even proof beyond a reasonable doubt, and the blameless may be held liable or the blameworthy excused. When this happens, the system has failed and we rightly feel unhappy. The use of explicitly statistical background evidence simply rubs our nose in this unpleasant fact. Daniel Shaviro, Statistical-Probability Evidence and the Appearance of Justice, 103 HARV. L. REV. 530 (1989). Sometimes the problem with background assumption evidence is said to be that it does not depend on the volitional conduct of the defendant. However, where liability is imposed erroneously, as it sometimes will be, the erroneously liable party will not typically have engaged in any conduct that seems volitional in the relevant sense. A person wrongly found liable or convicted based on erroneous eyewitness identification simply has the misfortune to look like the wrong person in the wrong place and time.

However, in a number of similar cases it would seem that juries would make background assumptions. Consider a case based on evidence that the offending bus was shocking pink and proof that the defendant owned a shocking pink bus. In this case the jury might well find liability based on an implicit assumption that shocking pink buses are rare, thus making by itself the background assumption which was regarded as problematic in the blue bus example.
instance, argued that non-veritistic concerns about both policy and verdict accuracy counsel against making explicit the inevitable uncertainty in verdicts, especially in the criminal setting.\textsuperscript{111} As Daniel Shaviro has noted, this policy argument assigns “a moral value to rationalization and denial.”\textsuperscript{112} Honesty about the possibility of error seems as likely to increase as to reduce verdict accuracy.\textsuperscript{113} Reminding jury members that they are fallible would seem to encourage a desirable humility and caution. The public generally is more likely to accept rules protecting defendants if they are acutely aware of the possibility of false conviction.\textsuperscript{114} In any event, concerns about quantification seem to have little relevance to evidence of societal discrimination patterns, which is typically relatively qualitative in nature (e.g. the level of discrimination is high or low).

The most troubling aspect of using background evidence is a problem that is seldom raised. In several recent articles, Chris Sanchirico has explored the incentive effects of evidentiary rules on the underlying behavior that the law seeks to deter.\textsuperscript{115} Incentive effects seem to explain a number of anomalous evidentiary rules, including the treatment of character evidence. An individual with a dubious past has no incentive to avoid future bad acts if the evidence of past bad acts will be admissible and will almost certainly lead to conviction, no matter what future choices the individual makes.\textsuperscript{116} Excluding character evidence would seem to reduce the risk of conviction and increase the individual’s incentive to avoid future bad acts.

\begin{footnotesize}
\begin{enumerate}
\item[111] Trial by Mathematics, \textit{supra} note\underline{___}.
\item[112] Statistical-Probability Evidence and the Appearance of Justice, \textit{supra} note\underline{___}, at 535.
\item[113] Id. at 547-48.
\item[114] Id. at 547-8.
\item[116] Id.
\end{enumerate}
\end{footnotesize}
evidence gives such a person the same incentive to behave as well as the average member of the community. Incentive considerations raise the concern that allowing accurate information about high societal levels of discrimination will discourage anti-discrimination efforts by firms who assume they will be found liable anyway because of high societal discrimination levels. Conversely, perhaps evidence of low societal levels of discrimination would encourage free-riding on the societal level of discrimination by those who lived in low discrimination areas. Unfortunately, there seems to be no way around this problem when jurors bring strong background assumptions into the courtroom. One way or another, their use of background assumptions may create undesired effects. Improving the accuracy of background assumptions through evidence does not address incentive problems, though it does not seem to worsen them either.

E. Reliability

Not all potential evidence is reliable, and in some instances the jury’s pre-existing views may be preferable. In the landmark Daubert v. Merrell Dow Pharmaceuticals,\(^ {117}\) the Supreme Court directed lower courts to act as gatekeepers for the reliability of expert evidence. Evidence about the societal rate of discrimination is and should be subjected to the Daubert requirements for reliability. These threshold tests are generally considered stringent and seem to have significantly increased the rate at which expert testimony is excluded on grounds of reliability.\(^ {118}\) Any concerns that the Daubert framework is too permissive should be dealt with by tightening those standards.


Yet social framework evidence about discrimination may seem to raise reliability issues that are beyond the usual Daubert analysis. Experts are no more exempt than the rest of us from the logic of Bayes’ Law. They choose their experimental designs and evaluate the experimental results of others in the light of their own background assumptions. Certainly empirical research on discrimination is influenced by a variety of factual priors with political overtones. Still, even a partisan research agenda can produce useful studies that are reliable in the sense required by evidence law. A significant and productive part of scientific inquiry involves not neutral hypothesis testing but efforts to defend core beliefs.119 Even if all experts stubbornly insulated their core beliefs from new evidence, those experts could play a useful role by synthesizing the empirical case for their own side. The task of synthesis requires specialized knowledge beyond the usual participants in the adversary system, and the strength of that system is providing a forum for each side to present its strongest arguments, in this case through expert testimony.

**F. Reference Class Problems**

Up until now I have spoken of the general societal pattern of discrimination. The pattern of discrimination, however, is in all likelihood not constant across a society.

---

Rather, it varies between various reference classes such as region and profession.120 This complication is at the heart of the so-called reference class problem, which is actually not a single problem but a series of related problems.

Reference class problems have vexed mathematicians, philosophers and evidence scholars for over a century.121 One important implication of these problems is that the concept of probability is unavoidably relative to a reference class, or, in other words, that all probability is conditional. There are numerous possible reference classes, none of


which is privileged.\textsuperscript{122} This still leaves the difficult question of how to choose reference classes for purposes of formulating background assumptions, but it has refocused the question in a useful way. Reference class problems are epistemological rather than metaphysical and can be addressed only in relation to the state of our knowledge in a given situation.

Not all reference class problems are equally serious. The simplest such problem occurs when we have information about only a single reference class, though one that is directly applicable to the case in question. For instance, suppose a plaintiff wishes to introduce evidence about the overall pattern of discrimination in the contemporary US business world. The defendant asserts that this pattern may not apply to its own operations, but presents no evidence at all to support this. Under these circumstances, the reference class problem does not cut very deep. We are always prisoners of our limited knowledge. Unless a defendant can provide more, the overall US pattern is helpful information.\textsuperscript{123}

Now suppose that the defendant does supply additional evidence that its own circumstances are not entirely typical, but that evidence is less reliable than the economy-wide data. In formal terms, there is more second order uncertainty about the accuracy of the more specific estimates. For example, suppose that in an individual discrimination case, the defendant demonstrates that its area and industry are more integrated than the economy as a whole. If the two pieces of evidence have equal second order uncertainty, the more specific evidence should supersede data on the general societal pattern. Whether this principle applies with different second order uncertainty is the subject of ongoing

\textsuperscript{122} A. Hájek, \textit{The reference class problem is your problem too}, 156 \textit{SYNTHESE} 563–585 (2007).

\textsuperscript{123} Veridical Verdicts, \textit{supra note} \textsuperscript{___}. 
debate. The reference class problem here is somewhat more serious, and may caution against applying Bayes’ Law in a formal manner. However, the implication of disputes about second order uncertainty is not that background evidence should be jettisoned, but that information about both reference classes is relevant, leaving the balancing of the two types of evidence to the practical judgment of the jury.

Another reference class problem arises when information is available for two distinct classes to which the defendant belongs. Suppose that the plaintiff is a dock worker, the kind of traditionally male occupation in which sexism is common, but that the workplace is located in a bastion of progressive thought such as San Francisco. Background evidence is available on discrimination in a nation-wide sample of ports and on all workplaces in San Francisco. Which information is more relevant? It might seem at first that the answer to this question must be based solely on data from a large group of San Francisco ports. In fact, no statistician would approach the problem by partitioning the data into reference classes but would rather use a large sample of varying locations and occupations and employ various statistical methods such as regression to make a prediction of the marginal contribution of each factor. Even data points that do not share many of an establishment’s characteristics can be useful in analyzing the establishment.

The most serious reference class problem arises only when data has been collected about each reference class separately. For example, suppose the EEOC conducted a study

______________________________


125 E. K. Cheng, A Practical Solution to the Reference Class Problem, 109 COLUM. L. REV. 2081 (2009). An alternative and inferior method is to use statistical methods to choose the more predictive reference class. Id.
of discrimination by location but did not collect occupational data. The NBER conducted a study of discrimination by occupational data but did not collect geographic data. Both studies provide helpful information but there is no way of using statistical methods to integrate the data they provide. The absence of algorithmic methods of aggregation does not, however, render the information irrelevant. It simply means that the integration of the information must be accomplished not formally but though the practical sense of the jury.

VI. INSTITUTIONAL CONCERNS

Evidence of the societal pattern of discrimination is relevant to any discrimination case in which the factual occurrence of differential treatment is disputed. Yet admitting social framework evidence of this pattern into virtually all discrimination trials would dramatically change the nature of those trials. Each discrimination case would necessarily become an open forum for disagreements about the degree to which discrimination remains a problem in society generally.

The problem would be especially acute in smaller individual cases. At present, social framework evidence is limited to a small set of class actions. These are often large, well-financed cases that pit opposing teams of well-known experts against each other. Individual litigants, however, can seldom afford the costs of trial by expert.

This Part will first examine the case for subjecting social framework evidence to appellate review. By creating a stable body of precedent, such review would avoid undue burdens on litigants in run-of-the-mill cases and avoid creating an enormous body of controversial and inconsistent decisions on a divisive social question. The Part will then examine mechanisms by which social framework evidence introduced through expert testimony can be made subject to review, and which of those mechanisms is best suited to discrimination law. Finally it will examine whether experts should be limited to pure
social framework evidence or whether they may also provide applied social framework evidence.

A. Reviewability

The foundational analysis of social science evidence was laid out in a series of papers by John Monahan and Laurens Walker. In non-legal settings, social science would clearly be classified as factual in nature: social science (we hope) describes how the world works, not how we would like it to work. However, social science facts differ in an important way from the canonical common law fact. The traditional common law envisioned facts as of interest only to the parties in the case at hand. In part for this reason, findings of fact are generally not reviewed on appeal, except in certain cases of clear error. As a corollary, they have no precedential value.

126 Walker and Monahan originally developed their framework to apply to social science, although they later extended it to other types of scientific inquiry. Laurens Walker & John Monahan, Scientific Authority: The Breast Implant Litigation and Beyond, 86 VA. L. REV. 801 (2000). In this later work they use the term “scientific authority” rather than “social authority,” presumably with corresponding changes to “scientific fact” and “scientific framework.” For present purposes, I will refer to all non-science evidence as traditional evidence, although there may be types of evidence that are in some other sense not traditional.


128 Civil law allows appellate courts broader review of facts.

129 In this discussion I stress the interests of other parties as an explanation for the different treatment of law and facts. Other rationales have been offered. The closeness of the trier of fact to the specifics of the case in some instances tends to make those triers more accurate than appellate
Legal conclusions, in contrast, may be binding on subsequent parties and are of importance to a wider audience than are facts. Typically, findings of law are fully reviewable on appeal. The deliberative processes of a given judge may exhibit individual quirks, and review by appellate courts subjects legal rulings to the scrutiny of many eyes, reducing the role of idiosyncratic judgments.\textsuperscript{130}

Like findings of law, social science may be at least somewhat general in its application, and the policy considerations surrounding the treatment of social science may be more like those surrounding law than like those concerning fact. Monahan and Walker classify social science in legal proceedings according to the generality of the social science and the use to which it is put.\textsuperscript{131} General social science introduced to establish a legislative fact relevant to a rule of law is social authority.\textsuperscript{132} Specific social science used to prove an adjudicative fact relevant to the case at hand is social fact.\textsuperscript{133}

courts. The appellate court is, for example, less able to assess witness demeanor. For matters of law, the value of proximity is slight or even nonexistent. Many commentators regard the comparative advantage of trial courts as the principal justification for limiting appellate review of facts. But though witness demeanor evidence may have been of great importance in the earlier times when the basic nonreviewability of facts was established, it is today a relatively exceptional type of evidence, and for the more typical documentary and other evidence, trial courts seem to have no clear advantage. Indeed, the advent of low cost video technology may reduce the advantage of trial courts in assessing demeanor.


\textsuperscript{131} This distinction was first made in Kenneth Culp Davis, \textit{An Approach to Problems of Evidence in the Administrative Process}, HARV. L. REV. 364 (1942).

\textsuperscript{132} Social Authority, \textit{supra} note___.

\textsuperscript{133} Social Facts, \textit{supra} note____.
General social science used as an aid in evaluating adjudicative evidence is social framework analysis. The core proposal of the Walker and Monahan plan is that social science should be treated like fact to the extent that it is fact-like and like law to the extent it is general, extending beyond the case at hand, and therefore law-like. A finding of social fact should have no precedential value, although determinations about the methodology by which social facts are established may have precedential value. For example, precedential value should be accorded to a finding that a particular forensic method or statistical technique is valid, but no precedential value should attach to the particular finding of fact resulting from the use of that technique. Findings about social frameworks and social authority should receive more judicial oversight and should correspondingly be accorded more precedential value.

B. Mechanisms of Review

In part to facilitate appellate review, Monahan and Walker originally proposed that law-like social science be presented to trial courts through legal briefs rather than expert testimony. Trial courts would then use that evidence reach factual conclusions. Factual conclusions drawn from evidence presented as social authority would be used to formulate a rule of law. Factual conclusions drawn from social framework evidence would be communicated to the jury though instructions. Both rules of law and jury

134 Social Frameworks, supra note ___.
135 Social Authority.
136 Social Frameworks at 583-596; Social Authority at 496-97
137 Expert testimony was to be used only to determine social facts, and even then only for non-methodological issues. The choice between briefs and expert witnesses is purely a matter of policy, since both routes are consistent with evidence law. Although the Federal Rules of Evidence contain no explicit guidance, the Advisory Committee Notes clearly indicate that the
instructions would be supported by analysis in the opinion of the relevant social science evidence. Appellate and future lower courts would therefore have explicit arguments and instructions that could be reviewed and refined. 138

In a recent article, Walker and Monahan, along with Gregory Mitchell, modified the proposed treatment of social frameworks. 139 They noted the widespread rejection of their plan to replace expert witnesses with briefs and jury instructions even by those supportive of the concept of social frameworks. 140 Though not rejecting the possibility of presentation through briefs and the use of jury instructions, they accepted the criticism of their earlier work and endorsed the use of expert witnesses and the absence of guidance through jury instructions. 141

The use of expert witnesses rather than briefs to communicate specialized information seems perfectly consistent with the goals of the original Walker-Monahan proposal, 142 but the abandonment of jury instructions is more troublesome. Walker and drafters endorsed virtually any route by which judges might gather information about non-adjudicative facts. FEDERAL RULES OF EVIDENCE 201, Advisory Committee Note a.

138 Social Authority at 512-16.

139 Contextual Evidence of Gender Discrimination, supra note__.

140 Contextual Evidence of Gender Discrimination, supra note__, at 1729-31

141 Contextual Evidence of Gender Discrimination, supra note__, at 1732.

142 Social science research cannot be understood piece by piece: each line of inquiry must be understood as a whole. Even honorable lawyers will inevitably cherry-pick among studies when writing briefs, and indeed it is their job as advocates to do so. In matters of law, where judges are expert, they can effectively integrate two opposing pieces of advocacy, filling in the missing pieces with their own knowledge and research. Judges cannot perform this function in evaluating social science – that is the job of expert witnesses, who can be questioned by both parties and also by the court. Although presentation through brief may on occasion be appropriate, expert
Monahan’s responsiveness to their critics is admirable, but I wish they had stood their ground a bit more with respect to the need for guidance by courts. At the heart of the Monahan-Walker proposals was the insight that review serves valuable functions for any type of general principle, whether normative or empirical.\textsuperscript{143} When expert testimony goes straight to the jury, the opportunity to create a reviewable finding is lost. Walker, Monahan and Mitchell suggest that judges may still comment more broadly on social framework analysis than on other evidence,\textsuperscript{144} yet such comments would be relatively toothless because they would lack precedential value.

Much of the discussion social framework literature has treated as mutually exclusive the option of allowing juries to hear expert testimony and that of using expert testimony to create jury instructions or other reviewable structure.\textsuperscript{145} Yet there is no obvious inconsistency between these two uses of expert evidence. An alternative to jury instructions is to use social framework evidence to create evidentiary devices,\textsuperscript{146} a role witnesses seem to be the preferable method for presenting social science both to judges to influence findings of law and to juries to provide context for findings of fact.

\textsuperscript{143} Social Authority at 499-500.

\textsuperscript{144} Contextual Evidence of Gender Discrimination, supra note___, at 1742.

\textsuperscript{145} Contextual Evidence of Gender Discrimination, supra note___ at 1729-31 (surveying literature on use of expert witnesses).

\textsuperscript{146} The evidentiary devices I have in mind have been described by a variety of terms including presumptions, burdens of proof and burdens of production. This terminology has been rightly criticized on a variety of grounds, see, e.g. Ronald J Allen, Presumptions, Inferences and Burden of Proof in Federal Civil Actions--An Anatomy of Unnecessary Ambiguity and a Proposal for Reform, 76 NW. U. L. REV. 892 (1981); Ronald J Allen, Structuring Jury Decisionmaking in Criminal Cases: A Unified Constitutional Approach To Evidentiary Devices, 94 HARV. L. REV.
that seems to fall somewhere between social authority and the original conception of social frameworks. Canonical social authority, like evidence of the harm caused by segregation, seeks to influence a substantive rule that directly governs behavior. Canonical social frameworks seek to influence how juries interpret facts. The proposed hybrid would use an evidentiary device to shape how juries interpret facts. The evidentiary device does not eliminate the background assumption but rather determines it at the level of the rule-maker, typically the legislature or the highest court. Some evidentiary devices have already been used in Title VII jurisprudence and will be explored at length later.\textsuperscript{147}

Embodying an empirical assumption in an evidentiary rule creates precedent for other courts in the same jurisdiction to follow, refine, or reject. Formulation of these evidentiary devices should be an evolutionary process. Litigants should always have the right to introduce new expert testimony to support review of the rule. This process is critical to the goal of achieving consistency without undue rigidity.\textsuperscript{148}

A difficult problem seems to arise when an empirical question is raised by a case but the social science evidence is either wholly lacking or inadequate to provide a complete

\begin{enumerate}
\item \textsuperscript{147} See infra Part __ (arguing that using social framework analysis as a basis for evidentiary rules can provide a level of reviewability that would add much needed predictability to discrimination law).
\item \textsuperscript{148} In the federal courts, this process functions reasonably well within each Circuit. However, the creation of a reviewable rule does not guarantee that the Supreme Court will actually exercise review in a meaningful way. See infra Part VII (suggesting that the Supreme Court has been derelict in fulfilling its duty to resolve conflicts among lower courts).
\end{enumerate}
answer. This problem, however, is not as conceptually distinct as it might first seem from
the basic problem posed by social science evidence, and Monahan and Walker propose
that the missing fact be analyzed by analogy to the treatment of the relevant evidence.\footnote{149}
A court devising a common-law rule that must rest on factual assumptions would ideally
use evidence categorized as social authority. When such authority is significantly
inadequate and the court must still provide some rule, the court’s opinion should still give
careful attention to the factual assumptions of the rule, describing any relevant
information and explaining its speculation about factual issues that cannot be determined
empirically.\footnote{150} Emphasizing the role played by empirical assumptions is critical, since it
allows the rule to be refined as better evidence comes along.\footnote{151}

Social framework problems without social framework evidence provide a slightly
different problem, since the court could simply leave the factual questions in the case to
the jury unaided by judicial structure. This is in fact Monahan and Walker’s proposal,\footnote{152}
yet analogy to their original plan for social framework evidence\footnote{153} suggests that some
guidance by the court may be in order. Most importantly, jury instructions can frame the
background assumption that the jury must make and emphasize its factual as opposed to
normative nature. The court may suggest how the empirical problem could be analyzed
and identify the subsidiary issues that might bear on this determination (such as the
institutional constraints within the firm that increase or reduce discrimination). The

\footnote{150} Empirical Questions, supra note __, at 578-82.
\footnote{151} Empirical Questions, supra note __, at 578-82.
\footnote{152} Empirical Questions, supra note __, at 591-93.
\footnote{153} Social Frameworks at 595-96.
critical difference between this and the canonical social framework situation is that the
court would not actually provide the jury with a factual background assumption to use.
Instead, jury instructions would both provide guidance to the jury, and perhaps even more
importantly, pave the way for the introduction of social framework evidence in
subsequent cases.

C. Discrimination Law

Evidence bearing on the societal rate of discrimination might thus be treated under
several different regimes. First, that evidence might be treated as social authority for an
evidentiary device that implicitly dictated a specified background assumption to the jury.
The court’s opinion would detail the evidentiary device, noting the background
assumption made, the strength of the evidence for its assumption, the role played by its
own speculation and any reservations it might have about the adequacy of the data to
support its conclusions. 154 Second, the court might reach a factual conclusion about the
background assumption and convey that assumption to the jury through instructions,
again providing in its opinion the rationale for its factual conclusion. 155 Finally, evidence
of societal discrimination rates might be given as social framework evidence to the jury,
perhaps with some judicial structure but without any firm finding about background

154 This is both the pure social authority treatment and the treatment proposed for
incompletely answered issues of social authority.

155 This is the approach envisioned by the original social framework analysis, Social
Frameworks, supra note __. The use of evidentiary devices to control background assumptions is
noted in Linda Hamilton Krieger, Burdens of Equality: Burdens of Proof and Presumptions in
assumptions imposed by the jury instructions. These approaches lie along a continuum. At one end, the social authority approach provides a greater role for the court as opposed to the jury and greater predictability but less flexibility. At the other end, the unconstrained social framework approach provides a greater role for the jury, more flexibility but less predictability.

In an influential and important article, Professor Deborah Malamud noted that evidentiary devices in discrimination cases tend to oversimplify the complexities of real case-specific facts. Certainly this is an important concern. It may be dispositive in cases in which background assumptions are relatively uncontroversial and different juries faced with similar facts are likely to reach consistent outcomes. But there are enormous divisions in contemporary society about the societal rate of discrimination. Without precedential constraints, triers of fact will be forced to rely heavily on their individual background assumptions, producing unpredictable and inconsistent determinations in factually similar cases. Evidentiary devices may oversimplify case specific facts, but they achieve a measure of the consistency needed for the rule of law. Numerous studies have examined whether individuals accept the results of judicial decisions which they

---

156 This is both the modified social framework approach, Contextual Evidence of Gender Discrimination, supra note ____, and the approach I have proposed for incompletely answered social framework evidence.

157 Last Minuet at 2232-34.

158 Professor Malamud is admirably candid in acknowledging that her (and my) background assumptions are not entirely consistent with public perception about the societal rate of discrimination. Last Minuet at 2260-61. My difference with her view on the desirability of evidentiary devices may be explained either by different views on the influence that these beliefs have on the determination of case-specific facts or by different views on the variance in the public perception, which I regard as quite high.
oppose on policy grounds, and whether such decisions reduce overall support for the judicial system. For the most part, disagreement with result of a decision has little effect on the perceived legitimacy of the decision or of the judicial system as long as people believe that result is determined by neutral and procedurally fair rules of law.\textsuperscript{159} Large variations in the outcome of cases because of differing background assumptions cannot be easily explained as the result of neutral rules, and thus threatens the perceived legitimacy of the legal system. This threat seems particularly acute where the background assumptions in question are highly controversial. In these cases, disagreement with the result does have some tendency to reduce perceived legitimacy,\textsuperscript{160} and further reduction because of any lack of consistency becomes a serious problem.\textsuperscript{161}


\textsuperscript{161} Special evidentiary rules seem especially common in areas like discrimination where background assumptions are controversial. These evidentiary rules embody a particular view on background assumptions and take some discretion away from triers of fact. For example, in rape cases in which consent is an issue, jurors must unavoidably consider the probability that a woman who had engaged in consensual sex would either lie or convince herself that she had not consented. That probability is obviously a profoundly controversial subject. In response, rape shield laws embody a presumption that there is little or no connection between past sexual history and consent. Some observers have argued that rape shield laws serve non-veridical goals. Prior
Using social framework evidence as the basis for evidentiary devices is complementary to the standard use of social framework evidence to guide juries, and the standard use can serve as a kind of laboratory. Lawyers learn from their own experience and that of other lawyers. The first attempt to introduce a particular type of social framework analysis at trial may not produce the clearest exposition by either side, and the time may not have come to embody the social framework in an evidentiary device. 162 Before devising an evidentiary device, courts might wish to wait until the social framework analysis is sufficiently developed but not longer; delay creates the problems of inconsistency that have plagued discrimination law.

One final note concerns the political dimension of presumptions. Prior to the rise of social framework evidence, the role of background assumptions received some attention sexual experience, they suggest, supports an inference of consent, but allowing such evidence would place unfair burdens on women’s sexual freedom. Though in complete support of this policy concern, I am unconvinced by the claim that past history is probative, at least in the way suggested. A complainant with a history of many partners has no obvious incentive to lie about whether she has had yet another, while a previously inexperienced complainant might be trying to protect a reputation for chastity. Evidence that a complainant had consented previously and not claimed she was raped seems to increase rather than reduce her credibility. This is a reference class problem. Among the reference class of all women, past consent is evidentiary of consent in the current instance. Among the reference class of rape complainants, past consent not accompanied by a rape charge is evidentiary of lack of consent in present case. The tendency of facts finders to choose the wrong reference class seems to be the soundest reason why past sexual history should be inadmissible.

162 See infra Part VII (arguing that the need for consistency in discrimination law is sufficiently compelling that the phase in which social framework is not given serious review should be relatively short).
from commentators (though not courts) in a few specific Title VII settings. These discussions tended to assume that any judicially prescribed presumptions would necessarily be more pro-plaintiff than the decisions made by trial courts and juries left to their own devices. I make no such assumption now, and accept the possibility that some evidentiary devices that might be adopted will not conform to my own relatively pro-plaintiff background assumptions.

163 For more detail see infra Part VII.A.
164 This assumptions is noted and analyzed in Last Minuet at 2260-61.
165 The strongest case for making presumptions systematically pro-plaintiff is given in an intriguing article by Linda Hamilton Krieger. Burdens of Equality, supra note ___. Krieger makes a distinction between transformative law, which seeks to transform social norms, and normal law, which reflects current norms. Transformative law is inherently countermajoritarian. Krieger’s fascinating paradigm is the efforts of the colonial British to eradicate practices such as bride burning in India. Transformative law can also arise in a federal system, and as Krieger dryly observes, “although the federal majority tends not to see it this way, federal law in this context functions much like colonial law.” When triers of fact are likely to resist the transformation, evidentiary devices may function to overcome such resistance, and they served such a function in the Indian case. At the time of its passage, Title VII surely constituted transformative law. Where it was most needed it was resisted in principle: the norm of non-discrimination itself was rejected. Evidentiary devices at that time served a defensible but clearly countermajoritarian function. At present, however, I believe that virtually no-one in the United States rejects the principle of non-discrimination, and Title VII is best treated as normal law. Misperceptions about the societal pattern of discrimination may be common, but they are best settled through majoritarian methods, and any role for evidentiary devices should be justified on grounds consistent with majoritarian principles.
D. Applied Social Frameworks

The use of expert testimony to establish social frameworks raises an issue that was avoided by the original Monahan and Walker proposal: may experts present applied social framework evidence, or should they be limited to presenting only pure social framework evidence? My principal focus in this Article is on the treatment of pure social framework evidence, but the treatment of pure and applied evidence raises some intersecting issues.

Monahan, Walker and Mitchell, along with others, have argued that experts should limit their testimony to their area of expertise and that experts should not be permitted to comment informally on the specific facts of the case.\textsuperscript{166} Circumscription of expert testimony, they argue, is based on the rules of evidence,\textsuperscript{167} policy concerns about the


\textsuperscript{167} The Federal Rules of Evidence provides that expert testimony is admissible only when
proper role of experts,\textsuperscript{168} and the rights, including constitutional rights, of opposing parties.\textsuperscript{169} Other commentators have suggested that evidence law permits experts to comment on case facts.\textsuperscript{170} They argue that excluding applied social framework evidence will essentially eliminate all social framework testimony of societal discrimination patterns\textsuperscript{171} and will unfairly disadvantage plaintiffs.\textsuperscript{172}

If, as I argue, appellate review of social framework evidence is an important goal, the case against applied social framework evidence is strengthened. If experts both testify as to social frameworks and express opinions about the case at hand, appellate courts may have difficulty disentangling the two strands, making appellate review far more difficult.

\begin{quote}
\textsuperscript{168} Contextual Evidence of Gender Discrimination, \textit{supra} note\
\textsuperscript{169} Julie Seaman persuasively argues that permitting experts to testify about the application of theory to facts poses serious risks to the right to trial by jury. Triangulating Testimonial Hearsay, \textit{supra} note\
\textsuperscript{170} Matter of Context, \textit{supra} note\
\textsuperscript{171} \textit{Id.}
\textsuperscript{172} \textit{Id.}
\end{quote}
At the same time, case law gives some support to concerns that excluding applied social framework evidence of discrimination will eliminate the use of any social framework evidence. Courts historically excluded as irrelevant expert testimony that was not tied to specific case facts.\textsuperscript{173} Courts increasingly accept the relevance of social frameworks,\textsuperscript{174} but relevance concerns do continue to surface.\textsuperscript{175} In subjective practice class actions, some courts have rejected evidence of the societal discrimination patterns.\textsuperscript{176} Concern that this evidence will be rejected as irrelevant probably contributes to the extremely limited use of this evidence in individual cases. Under these circumstances, one can hardly fault counsel and experts for attempting to connect their testimony to case specific facts. These problems may be best addressed by a clear rule of law providing that pure social framework evidence is relevant and generally admissible. Such a rule would end concerns that evidence of societal discrimination patterns must be justified by connecting that evidence to case facts.

\textbf{VII. DISPARATE TREATMENT CLAIMS}

\textsuperscript{173} See generally Social Frameworks, \textit{supra} note\_\_\_; Federal Rules of Evidence 402, \textit{supra} note\_\_\_; Federal Rules of Evidence 702, \textit{supra} note\_\_\_ (Expert testimony permitted if it will “assist the trier of fact to understand the evidence or to determine a fact in issue”).

\textsuperscript{174} Social Frameworks, \textit{supra} note\_\_\_.

\textsuperscript{175} See Chadwick v. Wellpoint, \textit{supra} note\_\_\_; Ray v. Miller Meester Advertising, \textit{supra} note\_\_\_. See also When Do Courts Think Base Rates Statistics Are Relevant?, \textit{supra} note\_\_\_ (general historical reluctance of courts to use background evidence data).

\textsuperscript{176} EEOC v. Morgan Stanley, \textit{supra} note\_\_\_ (excluded in part, admitted in part); EEOC v. Wal-mart, \textit{supra} note\_\_\_ (excluded as irrelevant); Puffer v. Allstate Ins. Co., \textit{supra} note\_\_\_ (excluded on grounds that general evidence failed to show connection to facts).
The core provision of Title VII prohibits employers from “discriminat[ing] against any individual . . . because of . . . race, color, religion, sex, or national origin.” The typical discrimination case turns on the purely factual question of whether the employer intentionally treated an employee or applicant differently “because of” membership in a protected class, and requires the court to determine whether the defendant’s actions were motivated by the purpose of treating the plaintiff differently because of a protected characteristic.

Many important doctrinal issues in employment law concern how differential treatment is proven. Underlying many disputes over evidentiary burdens are differing background assumptions about the societal pattern of discrimination. These judgments, however, are seldom articulated by courts and must be inferred from evidentiary doctrine along with occasional illuminating dicta. Current Title VII doctrine implicitly delegates to the trier of fact the role of generating background assumptions about the societal pattern of discrimination. This arrangement has produced indefensibly inconsistent outcomes in discrimination trials. Consistency cannot be restored unless evidentiary


178 Discrimination proof patterns distinguish between two theories of recovery: disparate treatment and disparate impact. Disparate treatment liability requires a showing of intent to discriminate while disparate impact does not. Both theories were originally derived by courts from the general prohibition on discrimination contained in the original Civil Rights Act of 1964 although this prohibition does not appear to include either a requirement of intent or a distinction between intentional and non-intentional cases. Civil Rights Act of 1964 § 703 (a), PL 88-352 88th Congress, H. R. 7152 as codified and amended 42 USC §2000e-2. The role of background assumptions in disparate impact cases is less clear, since the purpose of disparate impact is highly contested, and the importance of background assumptions depends on the justification for disparate impact liability.
doctrine embodies some background assumptions, empirically derived or otherwise, about the nature and frequency of societal discrimination.

A. Individual Disparate Treatment

Most discrimination suits are brought by individual plaintiffs. The case law governing individual claims has added layers of tangled evidentiary devices on top of the brief statutory language prohibiting discrimination. The treatment of background assumptions about the frequency of discrimination has undergone roughly two phases. Background assumptions have of course shaped the law in both but in neither one has evidence about background assumption played an important role. In the first, some courts assumed that discrimination was relatively common and developed evidentiary devices that embodied that assumption. In the second, the Supreme Court apparently concluded that the prevalence of discrimination was lower than implied by the evidentiary devices and eliminated many of them. Rather than devising new evidentiary frameworks, the

179 The core provision of the original Civil Rights Act of 1964 provides:

(a) Employer practices. It shall be an unlawful employment practice for an employer-

- 

(1) to fail or refuse to hire or to discharge any individual, or otherwise to discriminate against any individual with respect to his compensation, terms, conditions, or privileges of employment, because of such individual's race, color, religion, sex, or national origin; or

(2) to limit, segregate, or classify his employees or applicants for employment in any way which would deprive or tend to deprive any individual of employment opportunities or otherwise adversely affect his status as an employee, because of such individual's race, color, religion, sex, or national origin.

Court delegated enormous discretion to the trier of fact. This delegation has led to wildly inconsistent outcomes in discrimination cases.

1. The Basic Proof Pattern

The Supreme Court first provided the pattern of proof in individual disparate treatment cases in the 1973 case *McDonnell Douglas v. Green*¹⁸⁰ and clarified that proof pattern in 1981 in *Texas Department of Community Affairs v. Burdine.*¹⁸¹ In the prima facie case, a plaintiff was required to show that she belonged to a protected group; that she applied for and met the minimum qualifications for a job opening; that she was rejected; and that the position remained open. The burden then shifted to the employer to articulate a nondiscriminatory reason for the adverse action.¹⁸² The defendant did not bear the burden of proof of non-discrimination, but was only required to raise an issue of fact about whether it discriminated against plaintiff. If the employer succeeded, then in the third and final step of *McDonnell Douglas*, the burden shifted to plaintiff to show that the defendant acted from a discriminatory motive and that the petitioner's stated reason for respondent's rejection was in fact pretext."¹⁸³

*Burdine* held that the plaintiff could meet its burden “either directly by persuading the court that a discriminatory reason more likely motivated the employer or indirectly by showing that the employer's proffered explanation is unworthy of credence.” The method of so-called direct proof is simply the standard delegation to the jury of the job evaluating the evidence as it sees fit. From a Bayesian perspective, the jury is to use all admissible evidence to make factual findings; use these factual findings to infer a likelihood ratio;


¹⁸² *McDonnell Douglas*, *supra* note__.

¹⁸³ *McDonnell Douglas*, *supra* note__.
and then combine this likelihood ratio with its subjective view of the prior probability of discrimination to reach a conclusion about the posterior odds of discrimination. In contrast, the so-called indirect method of proof, or proof by pretext, was viewed by some Circuits as announcing an evidentiary device to guide the jury’s inferences. The jury was permitted to determine the truth or falsity of the defendant’s proffered reason. However, proof of falsity was as a matter of law sufficient to prove discrimination. This rule was based on two implicit premises: (1) an assessment of the likelihood ratio using the “relative plausibility” principle and (2) an empirical premise that societal level of discrimination was relatively high.

In the 1978 Supreme Court case, Furnco Construction Corp. v. Waters, the Supreme Court indicated a limited awareness that the McDonnell Douglas-Burdine

---


185 Linda Hamilton Krieger has noted the role of the pretext rule in constraining the use of juror background assumptions. The Content of Our Categories, supra note __, at 1226.

186 See, e.g. Juridical Proof and the Best Explanation, supra note __. The relative plausibility principle provides that the likelihood of any explanation can only be evaluated in comparison to the plausibility of other potential explanations.

approach was based on a background assumption that the societal rate of discrimination was relatively high. The Furnco Court stated:

A prima facie case under McDonnell Douglas raises an inference of discrimination only because we presume these acts, if otherwise unexplained, are more likely than not based on the consideration of impermissible factors . . .

This is an empirical rather than logical inference. A defendant’s failure to provide an alternative explanation might in theory occur because of discrimination, but also because of bad record-keeping, stupidity, or personality animosity. The Furnco Court indicated that its inference was based on an empirical assessment of the relative likelihood of various explanations. Unfortunately, it indicated no awareness that this inference was also based on an additional premise that the societal level of discrimination was relatively high.

After the 1989 Price Waterhouse v. Hopkins, a separate framework was applied to so-called mixed motive cases in which the defendant’s proffered explanation had an element of truth and was thus not “pretextual.” Because Justice O'Connor's Price

188 Id.


190 Price Waterhouse v. Hopkins, supra note__.

191 Price Waterhouse v. Hopkins, supra note__.
Waterhouse concurrence\textsuperscript{192} was clearer than the plurality,\textsuperscript{193} later courts looked to it rather than to the plurality for guidance. O’Connor proposed that all individual discrimination suits follow the first two steps of the McDonnell Douglas framework.\textsuperscript{194} At that point, the plaintiff could try to show “by direct evidence that an illegitimate criterion was a substantial factor in the decision.“\textsuperscript{195} If the plaintiff succeeded in making this showing, “the burden on the issue of causation [shifts] to the defendant.”\textsuperscript{196} If the plaintiff could not meet this test, then the case was to be decided under the general rules of McDonnell Douglas, with plaintiff bearing the burden of proof on both intent and causation.\textsuperscript{197}

Like the proof of pretext rule, the Price Waterhouse framework operated by embedding background assumptions in an evidentiary device. O’Connor identified direct evidence of discriminatory intent as especially strong evidence.\textsuperscript{198} The jury was permitted

\begin{itemize}
\item \textsuperscript{192} Price Waterhouse v. Hopkins (O'Connor, J., concurring), 490 U.S. 228 (1989).
\item \textsuperscript{193} The plurality opinion at times appeared to state, contrary to basic legal principles, that the term "because of" was not equivalent to a requirement of but-for causation.
\item \textsuperscript{194} Price Waterhouse v. Hopkins (O'Connor, J., concurring), 490 U.S. 278-79
\item \textsuperscript{195} Price Waterhouse v. Hopkins (O'Connor, J., concurring), 490 U.S. at 276. See also id. at 265, 277.
\item \textsuperscript{196} Price Waterhouse v. Hopkins (O'Connor, J., concurring), 490 U.S. at 276.
\item \textsuperscript{197} Price Waterhouse v. Hopkins (O'Connor, J., concurring), 490 U.S. 278-79.
\item \textsuperscript{198} Price Waterhouse v. Hopkins (O'Connor, J., concurring), 490 U.S. at 276.
\end{itemize}

Where an individual disparate treatment plaintiff has shown by a preponderance of the evidence that an illegitimate criterion was a substantial factor in an adverse employment decision, . . . a reasonable fact finder could conclude that absent further explanation, the employer's discriminatory motivation "caused" the employment decision.

to determine whether the direct evidence showed a substantial role for impermissible motives but if the jury found a substantial role, the task of inferring causation, unavoidably based on background assumptions, was partially taken from the jury by placing the burden of proof on the defendant.

2. The Decline of Evidentiary Devices

In 1993, St. Mary's Honor Center v. Hicks199 resolved the Circuit split on the significance of disproof of defendant’s proffered explanation. Hicks held that a finding of pretext was not always sufficient to compel judgment for plaintiff. The Hicks majority offered two reasons for rejecting in this context the Furnco presumption that, absent explanation, adverse action against an individual who meets the prima facie case was caused by differential treatment. First, although the Hicks defendant’s proffered reasons at trial turned out to be false, the Court concluded that the plaintiff had not ruled out the possibility of non-racial animosity.200 Moreover, Scalia expressed concern that an incorrect explanation might reflect not the absence of a legal explanation but inadequate record-keeping by defendants.201

The majority opinion by Justice Scalia and some commentators argued that Hicks simply extended prior law,202 while the dissent by Justice Souter and other commentators regarded Hicks as a departure from settled doctrine.203 Both Scalia and Souter focused

199 Hicks, supra note__.

200 Hicks, supra note__; Rise of the Personal Animosity Presumption, supra note__.

201 Hicks, supra note__.

202 Hicks, supra note__; Last Minuet, supra note __, atPIN.

203 St. Mary's Honor Center v. Hicks, (Souter, J. dissenting), 509 U.S. 502 (1993); Rise of the Personal Animosity Presumption, supra note__; Ann C. McGinley, Rethinking Civil Rights and
principally on analysis of previous opinions, and neither noted the critical policy rationale for their different results, a difference of opinion on the societal pattern of discrimination. If discrimination is common, an employer’s inability to provide an adequate explanation for its actions is highly probative. If discrimination is not common, failure to provide an explanation is far less probative since non-racial animosity or poor record-keeping become plausible alternative explanations.\textsuperscript{204}

Justice Souter did note another problem with \textit{Hicks}: it rejected a proof structure but replaced it with no other structure.\textsuperscript{205} Souter noted that the lack of evidentiary devices was likely to increase the time and cost of trials, burdening both the parties and judicial system. He might also have noted another problem: evidentiary devices that take some discretion away from the trier of fact reduce the ways in which variations in views about the societal pattern of discrimination could influence case outcomes. Those devices need not favor plaintiffs. The majority might reasonably have adopted evidentiary devices less friendly to plaintiffs but by failing to adopt any device made inconsistent determinations a virtual certainty.

Uncertainty about evidentiary standards was further increased when, two years after \textit{Price Waterhouse}, Congress passed the Civil Rights Act of 1991.\textsuperscript{206} In the following years, Courts of Appeal divided over how to integrate the 1991 Act, the O’Connor Employment at Will: Toward a Coherent National Discharge Policy, 57 OHIO ST. L.J. 1443 (1996); Calloway, \textit{supra} note ___. Of course, the frequency of non-discriminatory but unfair decision-making is also important, Last Minuet at 2255-57, since it affects the likelihood ratio.

\textsuperscript{204} Commentators did note this issue, Rise of the Personal Animosity Presumption, \textit{supra} note __; Calloway, \textit{supra} note ___. Of course, the frequency of non-discriminatory but unfair decision-making is also important, Last Minuet at 2255-57, since it affects the likelihood ratio.

\textsuperscript{205} St. Mary's Honor Center v. Hicks, (Souter, J. dissenting), 509 U.S. 502, 538 (1993).

opinion, and the basic principles of *McDonnell Douglas* as reinterpreted by *Hicks*. In 1999, Tristan Green noted that “the framework and methods of proof for proving disparate treatment under Title VII remain in a state of chaos.”207 One of the numerous points of contention among lower courts was whether evidentiary devices continued to play any role in proof.

The Court did nothing to clarify this confusion for nine years and then handed down an extremely narrow ruling, *Reeves v. Sanderson Plumbing Products*,208 which did little to reduce confusion in the lower courts since it left those courts with the freedom to make almost any evidentiary inference. In her concurrence in *Reeves*, Justice Ginsburg gently suggested to her colleagues that a clearer set of evidentiary devices might be desirable.209

The continued chaos after *Reeves* was summarized in 2002 by Judge McKeown, writing for the Ninth Circuit in *Desert Palace, Inc. v. Costa*:


209 The Court today holds that an employment discrimination plaintiff *may* survive judgment as a matter of law by submitting two categories of evidence: first, evidence establishing a "prima facie case" . . . and second, evidence from which a rational fact finder could conclude that the employer's proffered explanation for its actions was false . . . I write separately to note that it may be incumbent on the Court, in an appropriate case, to define more precisely the circumstances in which plaintiffs will be required to submit evidence beyond these two categories in order to survive a motion for judgment as a matter of law.
The... jurisprudence has been a quagmire that defies characterization despite the valiant efforts of various courts and commentators. Within circuits, and often within opinions, different approaches are conflated, mixing burden of persuasion with evidentiary standards, confusing burden of ultimate persuasion with the burden to establish an affirmative defense, and declining to acknowledge the role of circumstantial evidence.\textsuperscript{210}

The Supreme Court agreed to hear \textit{Desert Palace},\textsuperscript{211} and in 2003, a unanimous opinion written by Justice Thomas held that the plain language of the 1991 Act provided that any admissible evidence, not just direct evidence, could be used by plaintiffs to establish that an impermissible motivating factor played a role. \textit{Desert Palace} is, like \textit{Reeves}, in my view substantively preferable to prior law.\textsuperscript{212} Again however, the Court’s extreme reluctance to provide any evidentiary device left the law governing proof of discrimination at least as confused as it was before. The lower courts have split wildly on the implications of \textit{Desert Palace} and continue to express frustration with this state of affairs.\textsuperscript{213}

\begin{flushright}
\end{flushright}


\textsuperscript{211} Desert Palace, Inc. v. Costa, 539 U.S. 90 (2003).

\textsuperscript{212} I say this with greatest respect for Justice O’Connor, whose concurrence saved us from the unspeakable mess that might have been created by the plurality’s inability to grasp first year tort law.

\textsuperscript{213} See, e.g., Duffy v. Sodexho, 2008 U.S. Dist. LEXIS 93378 (2008); White v. Baxter Healthcare Corp., 533 F.3d 381 (2008). In some Circuits, \textit{Desert Palace} has actually made the situation more complicated. See the excellent summary in \textit{Id}. 
The Supreme Court continues to refuse to provide evidentiary devices. In 2006, the Court vacated an 11th Circuit opinion\(^\text{214}\) on the grounds that it applied an incorrect standard for determining when comparative evidence would suffice to prove pretext.\(^\text{215}\) In a brief \textit{per curiam} opinion, the Court merely noted the imprecision of the proposed test, stating, “This is not the occasion to define more precisely what standard should govern pretext claims based on superior qualifications.”\(^\text{216}\) If, fifteen years after \textit{Hicks}, a grant of certiorari to a case using the wrong standard is not an appropriate time to articulate the right standard, the Court simply does not want to set a standard. This refusal is an abdication of the Court’s responsibility to ensure the uniform application of the laws.

The volume of discrimination litigation is a significant burden on the federal courts. In 2006, employment discrimination cases, overwhelmingly individual, accounted for a staggering 22% of trials.\(^\text{217}\) Reducing litigant uncertainty about the outcome of trials should tend to reduce this volume, and this reduction would normally occur over time as courts settled issues of law, but cannot happen when so many important issues in discrimination law are purely a matter of fact.\(^\text{218}\)

\(^{216}\) Ash v. Tyson, \textit{supra} note \_, at 458.
\(^{218}\) Richard A. Posner, \textit{An Economic Approach to Legal Procedure and Judicial Administration}, 2 J. LEGAL STUDIES 399-458 (1973); Giuseppe Dari-Mattiacci & Bruno Deffains, \textit{Uncertainty of Law and the Legal Process}, 163 J. INSTITUTIONAL AND THEORETICAL ECONOMICS 627 (2007). The high volume may itself be in part a result of greater uncertainty. Discrimination cases have long gone to trial at higher rates than other federal cases, and the proportion is increasing. In 2006, employment discrimination cases were 22% of trials but only about 6% of all
A renewed search for evidentiary devices, however, cannot take place as long as both sides refuse to confront the fundamental issue that divides them, their varying background assumptions about the societal pattern of discrimination. Productive dialogue requires candor, and lack of candor leads unavoidably to stalemate, expressed as maddeningly opaque technical disputes.

Perhaps the most charitable explanation for the reluctance of Court to confront differences in background assumptions is a concern that no progress can be made in settling these differences. The admission of social framework evidence provides the beginning of a route out of this impasse. That evidence cannot yet definitively resolve the issues in dispute, but it can provide a starting point for the Justices to express as clearly as possible the nature of and the basis for their background assumptions. Openly acknowledging the empirical nature of the dispute may also have the salutary effect of inducing a measure of humility into judicial deliberation: reasonable minds will seldom differ on matters of logic but will often differ on matters of fact.219

3. The Doctrinal Setting for Reviving Evidentiary Devices

I have argued that the Supreme Court should revive the use of evidentiary devices in discrimination law, making explicit the background assumptions behind these evidentiary devices and preferably basing these background assumptions on social framework evidence. The doctrinal implementation of these changes, however, is complicated by the federal civil case terminations. One cause of low settlement rates may be uncertainty about the ultimate outcome resulting from the chaotic state of the law governing evidentiary burdens.

219 A related point is made in Psychological Realism, supra note __, at 26-27; Cultural Cognition at Work, supra note __, at 36-40; Whose Eyes, supra note __, at 897. The cultural cognition literature addresses the problems raised by bias in background assumptions but the same reasoning applied to rational differences, indeed perhaps even more strongly.
current state of the Title VII’s evidentiary scheme, which is routinely referred to as a “quagmire”, “swamp,” or “morass.” This Part therefore examines the doctrinal setting into which any new evidentiary device would be incorporated.

The current statutory scheme has three key provisions. The basic prohibition from the 1964 Act provides that it is an unlawful employment practice to discriminate against any individual “because of such individual's race, color, religion, sex, or national origin.”

The 1991 Act added the motivating factor rule which provides that a plaintiff establishes an unlawful employment practice by demonstrating that “race, color, religion, sex, or national origin was a motivating factor for any employment practice, even though other factors also motivated the practice.” The 1991 Act also added a limitation on damages when the defendant can prove the absence of causation.

When most of the major evidentiary cases were decided, the statutory framework contained only the basic prohibition, and it is useful to examine the statute in its current state before returning to those cases. Read together, these provisions are straightforward. The motivating factor rule is not a separate basis for liability but an elaboration of the basic prohibition of “unlawful employment practices.”

Plaintiff’s burden is to show


222 Civil Rights Act Of 1991 Section 107(a), codified in 20 USC § 2000e-2(m).


that an unlawful motivating factor played a role,\textsuperscript{225} and lack of causation is a limited defense to be proven by defendant.\textsuperscript{226} Had the statute read this way initially, no court would have attempted to bifurcate the treatment of mixed and single motive cases. Once Desert Palace clarified the obvious – that the motivating factor rule does not require direct evidence – the motivating factor rule must govern all individual disparate treatment cases. In theory, the motivating factor rule does not preclude other ways of proving an unlawful employment practice. However, any alternative scheme that began with the rule that the plaintiff has the burden of proving all elements of the case would be a dead letter that no plaintiff would chose to use.\textsuperscript{227}

Despite this substantially changed statutory framework, some lower courts have clung to the distinction between single and mixed motive cases.\textsuperscript{228} In light of the


\textsuperscript{226} Civil Rights Act Of 1991 Section 107(a), codified in 20 USC § 2000e-2(m).

\textsuperscript{227} As if the plain meaning of the statute were not enough, this unified treatment of mixed and single motive cases is supported by several policy considerations. To preclude plaintiffs who claim single motive from availing themselves of the motivating factor rule would have the absurd result of giving plaintiffs an easier burden when they conceded that the defendant’s motives were not entirely discriminatory. And as numerous scholars have pointed out, the distinction between single and mixed motive cases is premised on a simplistic view of human motivation. Content of Our Categories, supra note ___, at 1223.

\textsuperscript{228} The Fifth Circuit has essentially eliminated the distinction between mixed and single motive cases, permitting all plaintiffs in a mixed-motive case to rebut the defendant’s legitimate non-discriminatory reason either through evidence of pretext or through the mixed-motive alternative. Machinchick v. PB Power, Inc., 398 F.3d 345, 352 (2005); Rachid v. Jack in the Box, Inc., 376 F.3d 305, 312 (2004). The Eleventh Circuit has rejected extension of the motivating factor rule to single motive cases. Cooper v. Southern Co., 390 F.3d 695 (2004).
substantially changed statutory framework, it seems far more helpful to ask how, if at all, the insights of earlier cases might be adapted to the newer statutory framework. The increased evidentiary structure within the 1991 Act goes some way to reducing the role of jury background assumptions, but the Supreme Court may well wish to impose more structure yet.

4. Pretext Revisited

Prior to Hicks, courts that adopted the proof of pretext rule permitted plaintiffs to prevail if they could demonstrate that the defendant’s step-two proffered reason was

Circuits vary in the sharpness of the distinction they draw between the summary judgment standard and the jury instructions at trial. The Eighth Circuit and Eleventh Circuits have held that the holding of Desert Palace was restricted to jury instructions and “had no impact on prior . . . summary judgment decisions” so that McDonell Douglas/Burdine continues to govern the summary judgment analysis of mixed-motive claims. Griffith v. City of Des Moines, 387 F.3d 733, 736 (2004); Burstein v. Emtel, Inc, 137 Fed. Appx. 205, 209 n.8 (2005) (unpublished). The D.C., Fourth and Ninth Circuits permit a mixed-motive plaintiff to avoid a defendant's motion for summary judgment by proceeding either under the "pretext framework" or by the motivating factor test.” Diamond v. Colonial Life & Accident Ins. Co., 416 F.3d 310 (2005); McGinest v. GTE Serv. Corp., 360 F.3d 1103, 1122 (2004); Fogg v. Gonzales, 492 F.3d 447, 451 (2007). Desert Palace encouraged confusion in this matter by explicitly declining to address the applicability of the motivating factor test to single motive cases. Desert Palace, supra note___, at fn 1.

In the event that the Court does not follow this proposal, more incremental incorporation of empirical evidence is possible. Linda Hamilton Krieger has suggested that summary judgment rules offer a “soft entry point” for incorporating insights from expert testimony. Intuitive Psychologist, supra note___ at 45.
“pretextual.”230 The use of the word “pretext” was unfortunate, since it connotes conscious deceit and thus contributed to the confused distinction between mixed and single motive cases. Still, the basic reasoning of the pretext rule was sound. If a defendant cannot provide a legitimate reason that might have motivated its decision, then our estimate of the probability of an illegal reason surely rises.231 This reasoning applies with equal force to single and mixed motive cases. This confusion caused by the word “pretext” can be avoided by requiring plaintiff to prove “falsity” instead of “pretext.”

The consequences of proving falsity should depend on a frank assessment of the societal pattern of discrimination, preferably informed by evidence.232 Under the proof by pretext rule, plaintiffs met their burden of showing both motive and causation if they could show the falsity of the proffered reason.233 The 1991 Amendments offer a wider

230 McDonnell Douglas, supra note ___.
231 Burdens of Equality, supra note __, at 120.
232 One other consideration deserves some weight in assigning the significance to be accorded to evidence of falsity. Justice Scalia’s majority opinion in Hicks raised the concern that an employer might face difficulty in reconstructing the reason for its action at a later date and challenged the suggestion of Justice Souter, in dissent, that ordinary personnel records should suffice. In a purely descriptive sense, Scalia may have the better of this argument. Both academic research and the facts of many cases suggest that adequate recordkeeping and “standard personnel policies” are less common than Souter supposed. A rule that penalizes employers who cannot provide a truthful factor leading to the job action will create incentives for more recordkeeping, and the potential costs of this must be considered in formulating evidentiary rules. Though I myself believe that improved recordkeeping would not be costly and might even be beneficial to employers, Deborah M. Weiss, Workers as Financial Instruments, (MANUSCRIPT) (2011), this is clearly a matter wholly unrelated to evidence of background assumptions.
233 See discussion in Part I. The Basic Proof Pattern.
variety of consequences that might flow from proof of falsity. The proof by pretext rule might be replicated by making proof of falsity adequate to prove motivating factor and also to create an irrebuttable presumption of causality. Such a rule would be based on a background assumption that discrimination was relatively common, and perhaps that it had not changed much since the 1980’s when Burdine was handed down. An intermediate rule, based on a correspondingly intermediate assessment of the societal level of discrimination, would be to allow proof of falsity to meet the plaintiff’s burden of showing motivating factor while allowing defendants to show the lack of causality. If the Court concludes that the societal pattern of discrimination is low, it may reach the same conclusion as the Hicks Court did through ostensibly doctrinal reasoning. Proof that the defendant’s reason is false may be considered by the jury but would be insufficient by itself to warrant liability. Yet a Court that believes discrimination to be low should go further than Hicks and provide at least some guidelines as to how discrimination may be proven. Perhaps, as some courts have held, proof of pretext should be a necessary but not sufficient condition for liability. Though, at least absent further evidence, I would not personally support such a rule, my grounds would be my own views about the societal level of discrimination rather than a view that such a rule was inherently flawed or biased.

5. Similarly Situated Employees
Plaintiffs often try to establish that they were treated differently from similarly situated employees outside the protected class. Such comparisons are the indirect evidence most commonly accepted as probative of differential treatment.\textsuperscript{234}

Background assumptions play several important roles in reasoning about whether two employees are similarly situated. Consider a plaintiff who was disciplined for theft and claims that an employee outside the protected class was not disciplined for a similar act. The more similar the other employee, the less likely this discrepant treatment was non-discriminatory. If the other employee is truly identical to the plaintiff except being outside the protected class, the comparative evidence is strong and the posterior odds of discrimination seem high even if our background assumptions about the societal level of discrimination are low. If more differences between the workers potentially explain differential treatment, the comparative evidence seems weaker. If we believe societal discrimination levels are high, we may still conclude that discrimination is likely despite this weaker evidence, while if we believe societal discrimination levels are low, this relatively weak evidence will not convince us that discrimination has occurred.\textsuperscript{235}

The role of background assumptions plays out doctrinally in the stringency of the test for whether the plaintiff’s comparative evidence is adequate. Lower courts have applied a variety of tests to determine this adequacy. In the case of comparative qualifications evidence, some courts have held that plaintiff must prove that “no reasonable person . . . could have chosen the candidate selected over the plaintiff for the

\begin{footnotesize}
\begin{enumerate}
\item[\textsuperscript{235}] Why are Employment Discrimination Cases So Hard to Win, \textit{supra} note\textsuperscript{___}.
\end{enumerate}
\end{footnotesize}
job in question”\textsuperscript{236} while others have imposed a lower burden requiring proof that “a reasonable employer would have found a plaintiff to be significantly better qualified for a job.”\textsuperscript{237} Some have required that the two employees be “nearly identical” in all respects\textsuperscript{238} while others have emphasized that only “relevant”\textsuperscript{239} characteristics be “comparable.”\textsuperscript{240} Different standards could be applied in different settings. A lower standard might defeat a summary judgment motion by defendant while a higher standard might be applied in jury instructions. As noted earlier, the Supreme Court declined to clarify this matter recently in \textit{Ash v. Tyson}, issuing instead an extremely narrow per curiam opinion.\textsuperscript{241} Devising a standard for comparisons is an admittedly difficult job but is for that reason all the more urgent. Difficult determinations are the ones most likely to lead to inconsistent jury determinations and Circuit splits, and only a uniform standard from the Supreme Court can produce a regime of uniform application of the laws.

\textbf{B. Class Actions}

The basic proof pattern for class disparate treatment claims was laid down in 1977 by two cases, \textit{Teamsters v. United States},\textsuperscript{242} and \textit{Hazelwood School District v. United States}.\textsuperscript{243} In a systemic disparate treatment suit, the plaintiff must show that the defendant

\begin{itemize}
  \item \textsuperscript{236} Cooper v. Southern Co., \textit{supra} note__.
  \item \textsuperscript{237} Aka v. Washington Hospital Center, \textit{supra} note__.
  \item \textsuperscript{238} Rioux v. City of Atlanta, 520 F.3d 1269 (2008); Taylor v. Peerless Indus., 322 Fed. Appx. 355 (2009).
  \item \textsuperscript{239} Wright v. Murray Guard, 455 F.3d 702 (2006).
  \item \textsuperscript{241} Ash v. Tyson, \textit{supra} note__.
  \item \textsuperscript{242} Hazelwood School District v. United States, \textit{supra} note__.
  \item \textsuperscript{243} \textit{Id.}.
\end{itemize}
has a policy, pattern, or practice of discriminating against a protected group and that
discrimination is “the company's standard operating procedure, the regular rather than the
unusual practice.”244 Plaintiffs must provide statistical evidence that the practice leads to
differential outcomes for members and nonmembers of the protected group245 and must
supplement that analysis with non-statistical evidence.246

The basic structure of the statistical component of proof was described earlier.247 A
statistician calculates the likelihood that the employer’s workforce could have arisen
from chance given the workforce size and the proportion of the protected class in that
workforce and in the relevant labor market. Teamsters allowed comparison to the general
workforce,248 but Hazelwood made it clear that in most cases the employer’s workforce
must be compared249 to the “qualified …population” or relevant workforce.250 The

245 Statistical evidence is not an absolute requirement in systemic disparate treatment cases,
Pitre v. Western Elec. Co., 843 F.2d 1262 (10th Cir. 1988).
246 With extreme discrepancies, statistics alone may suffice, Hazelwood School District v.
United States, supra note ___, but such cases are very uncommon today.
247 Supra Part III.B.
248 Teamsters, supra note ___.
249 Such comparisons were permitted for unskilled jobs in the first case authorizing the use of
statistical evidence. Id.
250 In hiring cases, courts apply Hazelwood either by looking at the data on the relevant labor
market for the job in question, Hazelwood School District v. United States, supra note ____; Ward's
Cove Packing v. Atonio, 490 U.S. 642), see also EEOC Census 2000 Special EEO File,
http://www.eeoc.gov/stats/census/index.html (collecting relevant labor market data for use at trial)
or actual applicant data, reduced to include only those applicants deemed qualified. EEOC v.
Turtle Creek Mansion Corp., 70 FEP 899 (1995). See generally RAMONA L. PAETZOLD & STEVEN
Hazelwood Court indicated that it intended to leave further decisions about statistical evidence to trial courts.  

The only subsequent Supreme Court guidance on this subject occurred in the 1986 Bazemore v. Friday. The relevant workforce may be defined broadly or narrowly and a narrower definition provides a closer and more persuasive comparison. Unfortunately, the data available in discrimination litigation seldom if ever includes all variables that might be relevant. Bazemore held simply that even a statistical analysis that “includes less than ‘all measurable variables’ may serve to prove a plaintiff’s case” and that typically


It is thus clear that a determination of the appropriate comparative figures in this case will depend upon further evaluation by the trial court. As this Court admonished in Teamsters: "[S]tatistics . . . come in infinite variety. . . . [Their] usefulness depends on all of the surrounding facts and circumstances." . . . Only the trial court is in a position to make the appropriate determination after further findings.

Hazelwood School District v. United States, supra note__.

Bazemore v. Friday, 478 U.S. 385 (1986)(per curiam). Plaintiffs had introduced regressions that shown a significant effect for race controlling for education, tenure, and job title. The defendant had challenged these results on the grounds that they omitted relevant variables such as the differences in pay between employment sites.

Bazemore v. Friday, 478 U.S. 385 (1986)(per curiam)(opinion of Brennan, J., concurring) at 477.
“failure to include variables will affect the analysis' probativeness, not its admissibility.”

The details of implementing these principles were left entirely to trial courts: At what point does missing information render evidence inadmissible? How much does missing information reduce probative value? Should a defendant be able to attack a plaintiff's case by claiming the analysis should include variables for which the defendant has not provided coded data? Should plaintiffs be permitted to aggregate over different departments and job positions?

The proper treatment of missing information is inextricably tied to background assumptions about the societal pattern of discrimination. A study with more missing variables is weaker evidence, but courts that believe that societal discrimination levels are high may nonetheless conclude that discrimination is likely. They will thus be permissive about admissibility, assign greater weight to studies, and allow more aggregation. Though the reasoning is somewhat more complex, they are likely to disallow defendants to challenge studies for failing to include variables of which the defendant did not keep records. Those who believe societal discrimination levels are low will deem analyses with missing variables or a high level of aggregation to be inconclusive.

Court rulings on evidentiary statistical issues are strongly associated with rulings on several non-statistical issues, notably those that surround the use of subjective

\[\text{References}\]


255 A disaggregated analysis is in theory more accurate, but by choosing a small enough sample, even the most powerful real life connection can be made to appear statistically insignificant
evaluation. Courts have split widely on these issues. Some courts remain skeptical that subjectivity is a cognizable “practice” or that it meets the commonality requirement for class actions. Even if a company’s excessively subjective rules have led to discrimination, two individuals in different locations, jobs, or company divisions may have virtually no common chain of supervision, and it may be argued that they have not suffered from a single practice or suffered a common harm.

In theory, the statistical issue of whether class members have different job outcomes from nonmembers is almost entirely separate from whether a claim of subjectivity satisfies the “practice” and commonality requirements. A court might conclude that a particular subjective practice seldom leads to discriminatory outcomes while agreeing that a variety of employees, despite their differences, had been subject to essentially identical manifestations of a subjective process. Conversely, a court might regard a particular subjective practice as deeply suspect while believing that the harms suffered by employees were too distinct to be handled by a class action rather than individual claims.

256 For many years, most plaintiffs attempted to provide the non-statistical component of proof with evidence of discrimination against individual plaintiffs and evidence of discriminatory views among defendant’s decision-makers.

257 Watson v. Fort Worth Bank and Trust, supra note ___.

258 FEDERAL RULE OF CIVIL PROCEDURE 23(A).

259 One statistical issue may be relevant to the commonality and practice question. Perhaps subjective procedures should produce different effects at different levels of aggregation corresponding to the decision-maker alleged to be discriminatory. For example, if discrimination at the store manager level is alleged, the proper level of aggregation may be the store rather than, say, the geographic division. Daniel S. Klein, Bridging the Falcon Gap: Do Claims of Subjective Decisionmaking in Employment Discrimination Class Actions Satisfy the Rule 23 (a) Commonality and Typicality Requirements, 25 REV. LITIG. 131 (2006).
In practice, however, courts that readily accept subjectivity as a theory of liability generally give plaintiffs great latitude in defining the relevant population and in making other statistical comparisons. Butler v. Home Depot found for plaintiffs on the commonality issue and gave great weight to the plaintiffs’ study, which compared defendant’s workforce with retail establishments generally, rejecting a study by defendant that based comparisons on other hardware retailers. Other courts have rejected commonality claims and also found unpersuasive far more detailed statistical studies that control for all or virtually all available variables because they did not control for variables that were unavailable due to the defendant’s recordkeeping practices.


261 Cooper v. Southern Co., 205 F.R.D. 596 (2003), aff’d by Cooper v. Southern Co., supra note___; Puffer v. Allstate Ins. Co., supra note___. In Carpenter v. Boeing, the court affirmed summary judgment for defendants, rejecting a study that concluded that "[h]ourly female employees who are similarly situated to males with respect to job, grade, shift, department, and budget code are consistently and highly statistically significantly less likely to work overtime, to work less overtime, and to receive less overtime pay. This pattern is consistent across time." Defendant responded that the analysis should have contained more refined data on individual employees although such data were not available in electronic form. The Court agreed, stating that “Electronic data are undeniably more convenient, especially for use in statistical studies, but inconvenience does not excuse failure to collect the data.” Evidently the plaintiff was expected to hand-code the personnel files of a company with 163,000 employees. Carpenter v. Boeing, supra note___. See generally Bridging the Falcon Gap, supra note__.
Since these statistical and non-statistical issues seem wholly unrelated, the association between outcomes on the two suggests that court background assumptions on societal discrimination levels are in fact influencing both. To some extent the problem might be relieved if and when the law governing subjective employment practices is ever clarified. The use of background assumptions, however, is an unavoidable part of evaluating statistical evidence, and the Supreme Court’s “factual context”\textsuperscript{262} approach is not workable, since the factual context will unavoidably be seen through the lens of lower court background assumptions.

VIII. CONCLUSION

Discrimination law that is agnostic about the frequency of discrimination is simply not possible. Where the occurrence or non-occurrence of some event is at issue, an assumption about the overall frequency of similar events is not only helpful but psychologically and logically unavoidable, and will always be implicit in any conclusion about whether discrimination occurred.

Courts cannot be faulted for basing discrimination doctrine on their background assumptions about the frequency of discrimination. Less defensible is the Supreme Court’s apparent retreat from attempting to construct rules under which certain evidence can be found, as a matter of law, to imply certain consequences. Discrimination doctrine today implicitly delegates to most triers of fact the duty of assuming the societal pattern of discrimination. This delegation has led to hopelessly inconsistent determinations. This inconsistency makes it difficult for either employees or employers to know their rights and duties in advance without serving any other substantive goal. Inconsistency results not from sensitivity to context but from the brute luck of which judge hears a case.

\textsuperscript{262} Bazemore, \textit{supra} note \_, at 401.
Social framework evidence of the societal pattern of discrimination offers some hope for improving the consistency of judgments. Such evidence is neither irrelevant nor intrinsically prejudicial. However, social framework analysis of discrimination also raises special challenges. This evidence is almost unknown outside of class subjective practice cases, and its wider use would constitute a major change in the law. Relitigating the societal pattern of discrimination in every discrimination case would create enormously inefficient duplication of effort. Most importantly, the general nature of social frameworks suggests that appellate review of this evidence should be more extensive than that normally accorded to fact.

Devising mechanisms for appellate review of social frameworks has proven difficult, and perhaps no general approach is possible for all areas of law. A system of evidentiary devices could move discrimination law towards the best that can be hoped for, a system in which background assumptions are incorporated in a way that tends to produce consistency between adjudications.
IX. APPENDIX

This Part examines in more detail the explanation in Part III of the unavoidable use of background assumptions.

A. Bayes’ Law

Part III. B. sketched Bayes’ Law, the mathematical foundation for analyzing background assumptions. A more detailed explanation is provided here.

Call the hypothesis to be proven \( h \). The probability of the hypothesis being true in any randomly picked case is \( p(h) \), also called the unconditional probability of \( h \) or the prior probability of \( h \). Suppose that there is some type of adjudicative evidence \( e \) that tends to support an inference of hypothesis \( h \). The probability that \( h \), if true, will give rise to evidence \( e \) is called the conditional probability of \( e \) given \( h \), or \( p(e|h) \).\(^{263}\) We are ultimately interested in the probability of the hypothesis \( h \) given the evidence \( e \), \( p(h|e) \), or the posterior probability of \( h \). The values \( p(e|h) \) and \( p(h|e) \) are two different numbers, though they are often confused. However, Bayes’ Law shows a simple relation between them.\(^{264}\)

The version of Bayes’ Law most useful in legal settings is often the odds ratio formulation:\(^{265}\)

\[ \frac{p(e|h)}{p(h)} = \frac{p(e|h)}{p(h)} \]

In other words, \( p(e|h) \) is the percent of times that \( h \) is true that also produce evidence \( e \).

\(^{263}\) The conditional probability of \( e \) given \( h \) is defined as the probability that both \( e \) and \( h \) are true, \( p(e\&h) \), divided by the probability that the hypothesis is true, \( p(h) \).

\[ p(e|h) = \frac{p(e\&h)}{p(h)} \]

\(^{264}\) By the definition of conditional probability, \( p(e|h) = \frac{p(e\&h)}{p(h)} \) and \( p(h|e) = \frac{p(e|h)}{p(e)} \). We can combine these to eliminate \( p(e\&h) \) and produce the simplest version of Bayes’ Law,

\[ p(h|e) = p(e|h) \frac{p(h)}{p(e)} \]

\(^{265}\) Evaluating the simple version of Bayes’ Law for \( \neg h \)

\[ p(\neg h|e) = p(e|\neg h) \frac{p(\neg h)}{p(e)} \]
The terms \( \frac{p(h|e)}{p(-h|e)} \) and \( \frac{p(h)}{p(-h)} \) are called odds, or odds ratios. The term \( \frac{p(h)}{p(-h)} \) represents the prior odds, while the term \( \frac{p(h|e)}{p(-h|e)} \) shows the posterior odds. The numerator and denominator of an odds ratio must add up to one, since the numerator represents the probability that an event will occur and the denominator the probability that that event will not occur. The odds ratio itself may take any positive value. The term \( \frac{p(e|h)}{p(e|-h)} \) is called a likelihood ratio or Bayes factor. The numerator and denominator of the likelihood ratio can each range from 0 to 1 and their sum can range from 0 to 2. One implication of Bayes’ Law is that assumptions about \( p(h) \) are unavoidable in any inference from \( p(e|h) \) to \( p(h|e) \). Suppose that we were to decide that our estimate of \( p(h) \) was unreliable and thus should be ignored. We disregard the prior odds ratio and assume that the posterior odds is equal to the likelihood ratio. Bayes’ Law implies that this would amount to an assumption that the prior odds ratio is 1:

\[
\frac{p(h|e)}{p(-h|e)} = \frac{p(e|h)}{p(e|-h)} * 1
\]

and dividing this into Bayes’ Law evaluated at \( p(h|e) \) produces the likelihood ratio version of Bayes’ Law in the text.

\[266\] For this reason, \( p(h|e) \) can be easily calculated from the odds ratio. Let \( \frac{p(h|e)}{p(-h|e)} = k \).

\[
\begin{align*}
\frac{p(h|e)}{1 - p(h|e)} &= k \\
p(h|e) &= \frac{k}{1 + k}
\end{align*}
\]
A prior odds ratio of 1 in turn implies that $p(h)$ and $p(-h)$ are equal, and that discrimination and nondiscrimination are equally probable.

$$\frac{p(h|e)}{p(-h|e)} = \frac{p(e|h)}{p(e|-h)} \times \frac{.5}{.5}$$

We have simply substituted an assumption of equal probability for an empirical estimate.

**B. Class actions**

In a class discrimination case, $h$ is the hypothesis that members of protected class were treated differently from other employees, while $e$ is an observed discrepancy between the proportion of some protected group in the employer’s workforce and in the relevant labor market.$^{267}$

Readily available evidence makes it possible to calculate $p(e|-h)$, the probability of obtaining the observed discrepancy by nondiscriminatory chance. The requisite evidence $e$ is the number of employees in the firm;$^{268}$ the proportion of the protected group in the qualified population; and the proportion of the protected group in the firm. If the employer is not discriminating, then the chance that it will hire a member of the protected class in any given hiring decision is simply the proportion of the protected class in the relevant workforce, which equals the probability of a random draw from the workforce. If a nondiscriminating firm chooses a given number of employees from the relevant labor market, the binomial distribution can be used to calculate the likelihood that the resulting

---

$^{267}$ Where a job requires specific skills, the qualified population consists of workers with those skills. Few jobs are filled by nation-wide search, and the relevant labor market is also limited by some measure of proximity to the employer. Hazelwood School District v. United States, supra note ___.

$^{268}$ Most cases consider not the entire firm workforce but a particular job classification, but the basic principles are the same, and for expositional purposes I will speak of the employees in the whole firm.
workforce will have no more than the observed proportion of members of the protected class.\textsuperscript{269} The larger the firm, the less likely that nondiscriminatory procedures would produce a workforce with a lower proportion of protected class workers than in the relevant labor market.

The probability that discriminatory hiring would produce no more than the observed proportion of protected class workers, $p(e \mid h)$, is much more difficult to calculate than $p(e \mid \neg h)$.\textsuperscript{270} For expositional purposes we will assume that $p(e \mid h)$ is very close to 1,\textsuperscript{271} which is equivalent to assuming that the relevant labor market will always have a higher proportion of protected class members than the workforce of a discriminating employer of the defendant’s size. This assumption is more likely to be true when the number of employees in the firm is large or the degree of discrimination is high.

---

\textsuperscript{269} The binomial distribution is the discrete probability distribution of the number of successes $k$ in a sequence of $n$ independent yes/no experiments, each of which yields success with a known probability $p$. Each hiring decision is a single experiment. Let evidence $e$ be (1) the number of employees in the firm, $n$; (2) the proportion of the protected group in the qualified population, $q_m$; and (3) the proportion of the protected group in the firm, $q_e$. For a single hiring decision, let $c$ be true if the employer chooses a member of the protected class. If the employer is not discriminating, then the probability of $c$, $p(c)$, equals the probability of a random draw from the workforce, which is in turn equal to $q_m$. In other words, the hypothesis $\neg h$ that the employer is not discriminating is equivalent to the hypothesis that $p(c)=q_m$. The probability that a nondiscriminating firm would have no more than the observed number of protected group employees is obtained by evaluating the cumulative distribution function of the binomial distribution at $n$ trials, with probability $p=q_m$, and successes $k=q_e$.

The value of $p(c)$ bears no necessary relation to the value of $p(h)$. The probability of differential treatment $p(h)$ is the percentage of firms that discriminate (that is, for whom $p(c)<q_m$) while for any firm that does discriminate, $q_m-p(c)$ is a measure of the firm’s degree of discrimination. For example, high $p(h)$ and low $q_m-p(c)$ corresponds to pervasive discrimination across firms, but low levels of discrimination by the discriminators.

\textsuperscript{270} If the employer is discriminating against the protected class, then $p(c)<q_m$. The probability that discriminatory hiring would produce the observed proportion of protected class workers, $p(e \mid h)$, can be rewritten

$$p(e \mid h)=p(n, q_e, q_m \mid p(c)<q_m)$$

This number is much more difficult to calculate than $p(e \mid \neg h)$. Simple use of the binomial distribution requires that we specify a single value for $p(c)$, in other words, that we know just how discriminatory discriminators are. Calculating a probability for a range of values of $p(c)$ (all those less than $q_m$) is possible but the calculation is both complicated mathematically and requires a probability distribution for values of $p(c)$.

\textsuperscript{271} The fact that $p(e \mid h)$ might not equal 1 means that the current rule requiring a threshold level of $p(e \mid \neg h)$ is deficient in another way besides the fact that it ignores $p(h)$. 

A liability rule that makes use of statistical evidence must stipulate a value of the posterior odds ratio, $\frac{p(h|e)}{p(-h|e)}$, that meets some threshold requirement, such as making a prima facie case. One approach would be to require evidence of the likelihood ratio, $\frac{p(e|h)}{p(e|-h)}$, and to require evidence of, or make explicit background assumptions about, the prior odds ratio $\frac{p(h)}{p(-h)}$. Suppose however, that the legal system attempts to ignore the prior odds of discrimination, $p(h)$, perhaps because information about $p(h)$ seems unreliable. The liability threshold must therefore be calculated using only the likelihood ratio, or, if we consider it reasonable to assume that $p(e|h)=1$, using only the value of $p(e|-h)$. As we saw earlier, inferring the posterior odds ratio from the likelihood ratio amounts to an assumption that the prior odds ratio is 1, which requires an equal probability of discrimination and non-discrimination.

$$\frac{p(h|e)}{p(-h|e)} = \frac{p(e|h)}{p(e|-h)} \times \frac{.5}{.5}$$

The current legal standard requires that $p(e|-h)$, the chance that nondiscriminatory hiring produced the observed discrepancy, be no greater than .05. In a case that just meets that standard, the assumption that $p(e|h) = 1$ and $\frac{p(h)}{p(-h)} = 1$ implies posterior odds of discrimination equal to 20:

$$\frac{p(h|e)}{p(-h|e)} = \frac{1}{.05} \times \frac{.5}{.5} = 20$$

---

272 Like $p(e|h)$, $p(h)$ should ideally be calculated for the relevant labor market. In principle, the rate of discrimination might be uniform across labor markets even if the occupational distribution of protected class workers was not. On the other hand, discrimination might not be uniform. Some common labor market factors probably influence both $p(h)$ and $p(e|h)$. For example, more discrimination in a given occupation lowers the proportion of the protected class.
or a 95.2% probability of discrimination.\textsuperscript{273} If 10\% of firms in that market discriminate while 90\% do not, a likelihood ratio requirement of 20 is met when the probability of discrimination is about 69\%.\textsuperscript{274} If only 2\% of all firms discriminate, a likelihood ratio requirement of 20 is met when the probability of discrimination is about 29\%.\textsuperscript{275}

\textit{C. Informal Evidence: Individual Disparate Treatment}

Consider an employee discharged for stealing. Evidence $e$ is the plaintiff’s demonstration that another employee was not discharged for an identical theft, and the defendant’s proof that the other employee has a marginally better performance record. These facts together are a set of evidence $e$. Suppose that we estimate that the likelihood ratio is 20. Suppose the plaintiff in question is a black Southern factory worker in 1964. Even if discrimination played a role in only 10\% of all discharges of black employees, a likelihood ratio of 20 would mean that discrimination is more than twice as likely as non-discrimination.

\[
\frac{p(h|e)}{p(-h|e)} = 20 \times \frac{.1}{.9} = 2.2
\]

Assuming that race discrimination played a role in one in a thousand discharges of white employees at the time and place seems generous to plaintiffs. With that background

\textsuperscript{273} $p(h|e) = \frac{20}{1+20} = .952$

\textsuperscript{274} $\frac{p(h|e)}{p(-h|e)} = 20 \times \frac{.1}{.9} = 2.2$

\textsuperscript{275} $p(h|e) = \frac{22}{1+22} = .69$

$\frac{p(h|e)}{p(-h|e)} = 20 \times \frac{.22}{.98} = .41$

$p(h|e) = \frac{.41}{1+41} = .29$
assumption, the likelihood ratio of 20 implies that non-discrimination is fifty times more likely than discrimination.

\[
\frac{p(h|e)}{p(-h|e)} = 20 \times \frac{.001}{.999} = .0196
\]